Non Rigid Registration of Diffusion Tensor Images

Olivier Faugeras, Christophe Lenglet, Théodore Papadopoulo, Rachid Deriche

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Non Rigid Registration of Diffusion Tensor Images

Oliver Faugeras — Christophe Lenglet — Théodore Papadopoulo — Rachid Deriche

N° 6104

January 22, 2007

Thème BIO

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6104--FR+ENG

Non Rigid Registration of Diffusion Tensor Images

Oliver Faugeras* , Christophe Lenglet† , Théodore Papadopoulo‡ ,
Rachid Deriche§

Thème BIO — Systèmes biologiques
Projet Odyssée

Rapport de recherche n° 6104 — January 22, 2007 —39 pages

Abstract: We propose a novel variational framework for the dense non-rigid registration of Diffusion Tensor Images (DTI). Our approach relies on the differential geometrical properties of the Riemannian manifold of multivariate normal distributions endowed with the metric derived from the Fisher information matrix. The availability of closed form expressions for the geodesics and the Christoffel symbols allows us to define statistical quantities and to perform the parallel transport of tangent vectors in this space. We propose a matching energy that aims to minimize the difference in the local statistical content (means and covariance matrices) of two DT images through a gradient descent procedure. The result of the algorithm is a dense vector field that can be used to wrap the source image into the target image. This article is essentially a mathematical study of the registration problem. Some numerical experiments are provided as a proof of concept.

Key-words: registration, non rigid registration, partial differential equations, differential geometry, parallel transport, statistics, linear elasticity

* INRIA, 2004 route des lucioles, 06902 Sophia-Antipolis, FRANCE
† Siemens Corporate Research, 755 College Road East, Princeton, NJ 08540, USA
‡ INRIA, 2004 route des lucioles, 06902 Sophia-Antipolis, FRANCE
§ INRIA, 2004 route des lucioles, 06902 Sophia-Antipolis, FRANCE

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Recalage non-rigide d'IRM du tenseur de diffusion

Résumé : Nous proposons un nouveau cadre variationnel pour le recalage dense non-rigide d'IRM du Tenseur de Diffusion (IRM-TD). Notre approche repose sur les propriétés géométriques de la variété Riemannienne des distributions normales multivariées, équipée d'une métrique dérivée de la matrice d'information de Fisher. L'existence de formes closes pour les géodésiques et les symboles de Christoffel nous permet de définire certaines statistiques et de réaliser le transport parallèle de vecteurs tangents dans cet espace. Nous proposons une énergie, pour notre problème de recalage, dont l'objectif est de minimiser les différences entre statistiques locales (moyennes et matrices de covariance) de deux IRM-TD à travers une descente de gradient. Le résultat de l'algorithme est un champs de vecteurs dense qui peut être utiliser pour transformer l'image source vers l'image cible. Cet article est essentiellement une étude mathématique du problème de recalage. Des experiences numériques sont présentées dans le but d'illustrer la faisabilité de la méthode.

Mots-clés : recalage, recalage non-rigide, équations aux dérivées partielles, géométrie différentielle, transport parallèle, statistiques, élasticité linéaire

Contents

1 Introduction

We deal with the problem of estimating the geometric deformations between two diffusion tensor images. This is reminiscent of the problem of estimating the deformation of two images where the values at each voxel are real numbers [12]. This is solved by minimizing with respect to the deformation field h an error criterion that takes into account two sources of a priori knowledge:

1. The properties of the images intensities characterizing their similarity.
2. The constraints on the possible geometric deformations.

In our case the "intensities" are diffusion tensors. The problem of measuring their similarity is much more complicated and the corresponding gradient descent scheme becomes significantly more involved.

Previous works on the subject was initiated by Alexander *et al.* [1] by extending multiresolution registration techniques to DTI after having introduced various possible dissimilarity measures for such images [2]. In [30] and [29] the authors proposed to register three-dimensional scalar, vector and tensor data by matching areas with a high degree of structure and then interpolating the sparse estimated displacement field to the complete dataset. Other approaches like [18], [15], [25] and [28] rely on one or several transformation invariant tensor characteristics like the eigenvalues, the anisotropy measures, the apparent diffusion coefficient or even the tensor components to perform the registration. When several characteristics are used, which is often the case, multiple input channel registration methods like the demons algorithm [16] are used. In [35, 33] and then [34], the authors proposed a piecewise affine registration technique based on the L^2 inner product of diffusion profiles. They also investigate the tensors reorientation issue raised by Alexander *et al.* in [3]. Recently, Cao *et al.* [8] proposed to apply the framework of the large deformations diffeomorphic metric mapping to DTI. Finally, Leemans [19] introduced an affine multi-channel registration technique based on the mutual information as well as an original feature based registration method based on the curvature and torsion of fibers pathways. We also want to point out that a few recent works have used the Riemannian or Log-Euclidean metrics to characterize the properties of deformation fields [4, 6, 26] obtained by scalar images registration algorithms.

Contributions of this paper:

In this paper, we extend the approach presented in [12] to matrix-valued images $I : \Omega \to S^+(3)$. To our knowledge, this is the very first work to make use of the Riemannian structure of $S^+(3)$, proved to be relevant for DTI processing for instance in [27, 13, 5, 20], in a non-rigid DTI registration algorithm. The numerical implementation of the method is very tedious. We will illustrate the feasibility of the approach on two-dimensional synthetic datasets.

Organization of this paper:

We first set up the registration problem in section 2 and recall some important notions on $S^+(3)$. We then detail the regularization (section 3) and the data (section 4) terms of the initial value problem 3. We detail the computation of the gradient of the data term in section 5. Finally, we present numerical experiments in section 6.

2 The Registration Problem

We consider the problem of estimating the geometric deformations between two Diffusion Tensor Images (DTI). At a conceptual level, DT images are integrable bounded functions defined in $\mathbb{R}^n$, $n = 2, 3$ with values in $S^+(3)$ (noted S^+ in the sequel). As briefly recalled below, this space has a natural Riemannian structure. Bounded means that all observed diffusion tensors are within a ball of center $\mathbb{I}$, the 3×3 identity matrix, for the distance defined by equation 4. The same equation shows that the eigenvalues must lie between two strictly positive constants and therefore that the set of bounded diffusion tensors (for the Riemannian metric) is also bounded for the 2-norm and therefore for all the usual p-norms and the Frobenius norm.

These abstract images are not directly observable because of the physics of acquisition. What we call an image is an element of $C^\infty(\mathbb{R}^n, S^+)$, the space of infinitely differentiable functions. They are bounded and Lipschitz continuous as well as all their derivatives.

2.1 Statement of the problem

Let I_1 and I_2 be two images and $h : \Omega \to \mathbb{R}^n$ a vector field defined on a bounded and regular region of interest $\Omega \subset \mathbb{R}^n$. The registration or matching problem may be defined as that of finding a vector field h^* minimizing an error criterion between I_1 and $I_2 \circ h$. The search for this function is done within a set $\mathcal{F}$ of admissible functions so as to minimize an energy functional $\mathcal{I} : \mathcal{F} \to \mathbb{R}^+$ of the form

$$\mathcal{I}(h) = \mathcal{J}(h) + \mathcal{R}(h).$$

The term $\mathcal{J}$ is designed to measure the "dissimilarity" between the reference image I_1 and the h-warped image, noted $\mathcal{T}_h(I_2)$. We have the following proposition

Proposition 2.1.1. *If the relation between the two images I_1 and I_2 is a change of coordinates $x' = h(x)$ then the value $I_1(x)$ should be equal to the value $\mathcal{T}_h(I_2)(x)$, where*

$$\mathcal{T}_h(I_2) = Dh^{-1} I_2(h) Dh^{-T}. \tag{1}$$

Proof. $I_1(x)$ is a twice contravariant tensor. In the new coordinate system defined by $x' = h(x)$ it is equal to

$$I_1'(x') = Dh(x) I_1(x) Dh^T(x),$$

because of the way tensor components change with changes of coordinates. This new tensor should be equal to $I_2(x')$ and this yields the expression for $\mathcal{T}_h(I_2)(x)$. □

Note that other possibilities for $\mathcal{T}_h(I_2)$ have been considered in the literature (see [3] for instance). R. Sierra [31], has considered the case where one wants to preserve the determinant of I_2; this leads to

$$\mathcal{T}_h(I_2) = (\det(Dh))^{2/3} Dh^{-1} I_2(h) Dh^{-T}$$

In the following we consider that $\mathcal{T}_h(I_2)$ is defined by equation 1. The term $\mathcal{R}(h)$ is designed to penalize fast variations of the function h. It is a regularization term introducing an a priori preference for smoothly varying functions. Our error criterion is classically the sum of a data term $\mathcal{J}$ and a regularization term $\mathcal{R}$.

The set $\mathcal{F}$ is a dense linear subspace of a Hilbert space H whose scalar product is denoted by $(\cdot, \cdot)_H$. If $\mathcal{I}$ is sufficiently regular, its first variation (also called the Gâteaux derivative) at $h \in \mathcal{F}$ is defined (see, e.g., [10]) as

$$\delta_k \mathcal{I}(h) = \lim_{\varepsilon \to 0} \frac{\mathcal{I}(h + \varepsilon k) - \mathcal{I}(h)}{\varepsilon} \tag{2}$$

If the mapping $k \to \delta_k \mathcal{I}(h)$ is linear and continuous, the Riesz representation theorem [11] guarantees the existence of a unique vector, denoted by $\nabla_H \mathcal{I}(h)$, and called the gradient of $\mathcal{I}$, which satisfies the equality

$$\delta_k \mathcal{I}(h) = (\nabla_H \mathcal{I}(h), k)_H,$$

for every $k \in H$. The gradient depends on the choice of the scalar product $(\cdot, \cdot)_H$ though, a fact which explains our notation. If a minimizer h^* of $\mathcal{I}$ exists, then the set of equations $\delta_k \mathcal{I}(h^*) = 0$ must hold for every $k \in H$, which is equivalent to $\nabla_H \mathcal{I}(h^*) = 0$.
These equations are called the Euler-Lagrange equations associated with the energy functional $\mathcal{I}$. They give necessary conditions for the existence of a minimizer but they are not sufficient since they only guarantee the existence of a critical point of the functional $\mathcal{I}$. These critical points can be found in many ways, including methods for nonlinear equations. Rather than solving them directly the search for a minimizer of $\mathcal{I}$ is done using a gradient descent strategy. Given an initial estimate $h_0 \in \mathcal{F}$, a time-dependent differentiable function (also denoted by h) from the interval $[0, +\infty[$ into H is computed as the solution of the following initial value problem:

$$\begin{cases} \dfrac{dh}{dt} = -\Big(\nabla_H \mathcal{J}(h) + \nabla_H \mathcal{R}(h)\Big), \\ h(0)(\cdot) = h_0(\cdot). \end{cases} \tag{3}$$

The asymptotic state (i.e. when $t \to \infty$) of $h(t)$ is then chosen as the solution of the matching problem, provided that $h(t) \in \mathcal{F} \quad \forall t > 0$.

2.2 Precisions on the Riemannian structure of $S^+(n)$

In this section, we remind some basic concepts that will be useful for the following. We recall that $S^+(n)$ denotes the set of $n \times n$ real symmetric positive definite matrices, Σ. It is a subset of $M_n(\mathbb{R})$, the set of $n \times n$ real matrices. It is also a m_n-dimensional C^∞ submanifold of $\mathbb{R}^{m_n}$ ($m_n = n(n+1)/2$) whose local coordinates can be chosen as the m_n algebraically independent components of the elements of Σ. We note $\varphi_n : S^+(n) \to \mathbb{R}^{m_n}$ the natural coordinates mapping of this manifold. We recall that $T_\Sigma S^+(n)$, the tangent space at Σ of $S^+(n)$, coincides with the set $S(n)$ of $n \times n$ real symmetric matrices. This is a vector space which can be identified with $\mathbb{R}^{m_n}$ through the mapping φ_n. Elements in that space are contravariant vectors. We finally denote by $T^*_\Sigma S^+(n)$ the cotangent space at Σ of $S^+(n)$, the dual space of $T_\Sigma S^+(n)$. Elements in that space are covariant vectors. The basis of $T_\Sigma S^+(n)$ and $T^*_\Sigma S^+(n)$ are taken to be as in [20].

We recall the following theorem, see, e.g. [22]:

Theorem 2.2.1. *Let $\mathcal{E}$ be the set of real $n \times n$ matrices such that all the eigenvalues λ_i are such that $|Im(\lambda_i)| < \pi$. The restriction to $\mathcal{E}$ of the exponential is a diffeomorphism between $\mathcal{E}$ and $\exp \mathcal{E}$.*

There are two consequences of this theorem that are used in the sequel. The first one is the

Corollary 2.2.1. *The exponential is a diffeomorphism between $S(n)$ and $S^+(n)$.*

In other words, the exponential of any symmetric matrix is a positive definite symmetric matrix and the inverse of the exponential (i.e. the principal logarithm) of any positive definite symmetric matrix is a symmetric matrix. Moreover, both the exponential and the logarithm are continuously differentiable in $S(n)$ and $S^+(n)$, respectively.
The second one is the

Corollary 2.2.2. *The logarithm of a matrix with positive eigenvalues exists, and is unique and differentiable.*

Proof. Any such matrix belongs to $\exp \mathcal{E}$ defined in theorem 2.2.1. Therefore its logarithm exists, is unique and differentiable. $\square$

We introduce two notations

Definition 2.2.1. *We note* $\exp$ *and* $\log$ *the exponential and its inverse, the logarithm. Given $M \in M_n(\mathbb{R})$, we note* $\operatorname{dexp}(M, X)$ *the derivative of* $\exp$ *at M, applied to the element $X \in M_n(\mathbb{R})$. This is also sometimes called the derivative of the function* $\exp$ *at M in the direction X. In a similar manner, given $M \in \exp \mathcal{E}$ we note* $\operatorname{dlog}(M, X)$ *the derivative of the function* $\log$ *at M in the direction X.*

Details on the directional derivative of the matrix exponential and its computation can be found in [24]. However, to our knowledge, there is no previous work on the computation of the directional derivative of the matrix logarithm. As we will see in section 5, this will be a key component of our method. In [21], we proposed a novel formulation for the directional derivative of the matrix logarithm $\operatorname{dlog}(M, X)$ based on the spectral decomposition of M. We will show that it is in fact a linear function of its second argument X.

The geodesic distance between two elements Σ_1 and Σ_2 of $S^+(n)$ was described in [20] and is defined by

$$\mathcal{D}(\Sigma_1, \Sigma_2) = \sqrt{\frac{1}{2}\operatorname{tr}\left(\log^2\left(\Sigma_1^{-1/2}\Sigma_2\Sigma_1^{-1/2}\right)\right)}, \tag{4}$$

It is justified by corollary 2.2.1. At each point Σ of S^+, the metric tensor G acts on pairs of tangent vectors of $T_\Sigma S^+$ and defines an inner product. Its inverse G^{-1} is twice contravariant. For any real differentiable function f defined on S^+, one defines its differential, noted $Df = [\frac{\partial f}{\partial \mathbf{x}^1}, \cdots, \frac{\partial f}{\partial \mathbf{x}^{mn}}]$, with respect to the coordinates defined by the chart φ_n, a covariant vector, and its gradient, noted ∇f, which is a vector of $T_\Sigma S^+$. The relation between Df and ∇f is through the metric tensor:

$$\nabla f = G^{-1} D f \tag{5}$$

Equation 4 defines a real function on $S^+(n) \times S^+(n)$ which is differentiable. The gradient of $\mathcal{D}^2$ with respect to Σ_1, noted ∇_{Σ_1}, at Σ_1 and for some fixed Σ_2, is equal to [23]:

$$\nabla_{\Sigma_1} \mathcal{D}^2(\Sigma_1, \Sigma_2) = \Sigma_1 \log\left(\Sigma_2^{-1}\Sigma_1\right). \tag{6}$$

It is a vector of $T_{\Sigma_1} S^+(n)$, hence a symmetric matrix. This can also be seen from the general relation

$$\log(A^{-1}BA) = A^{-1}(\log B)A, \tag{7}$$

by writing

$$\Sigma_1 \log\left(\Sigma_2^{-1}\Sigma_1\right) = \Sigma_1 \log\left(\Sigma_1^{-1}\Sigma_1\Sigma_2^{-1}\Sigma_1\right) = \log\left(\Sigma_1\Sigma_2^{-1}\right)\Sigma_1 = \left(\Sigma_1 \log\left(\Sigma_2^{-1}\Sigma_1\right)\right)^T$$

It is tangent at Σ_1 to the (unique) geodesic between Σ_1 and Σ_2. In the following, we use the cases $n = 3$ and $n = 6$. To facilitate the reading of the formulas, indexes running from 1 to 3 are lower case Latin characters, e.g., $i = 1, 2, 3$, indexes running from 1 to 6 are upper case Latin characters, e.g., $I = 1, \ldots, 6$, and indexes running from 1 to 9 are lower case Greek characters, e.g., $\kappa = 1, \ldots, 9$.

3 Regularization term

This section studies the regularization part of the initial value problem 3, i.e. the term $\nabla_H \mathcal{R}(h)$. We choose concrete functional spaces $\mathcal{F}$ and H and specify the domain of the regularization operators.

3.1 Function spaces and boundary conditions

We begin by a brief description of the function spaces that will be appropriate for our purposes. In doing this, we will make reference to Sobolev spaces, denoted by $W^{k,p}(\Omega)$. We refer to the books of Evans [11] and Brezis [7] for formal definitions and in-depth studies of the properties of these functional spaces.

For the definition of $\nabla_H \mathcal{I}$, we use the Hilbert space

$$H = \mathbf{L}^2(\Omega) = \underbrace{L^2(\Omega) \times \cdots \times L^2(\Omega)}_{n \text{ terms}} = \left(W^{0,2}(\Omega)\right)^n.$$

The regularization functionals that we consider are of the form

$$\mathcal{R}(h) = \kappa \int_\Omega \varphi(Dh(x))\, dx, \tag{8}$$

where $Dh(x)$ is the Jacobian of h at x, φ is a quadratic form of the elements of the matrix $Dh(x)$ and $\kappa > 0$. Therefore the set of admissible functions $\mathcal{F}$ will be contained in the space

$$\mathbf{H}^1(\Omega) = \left(W^{1,2}(\Omega)\right)^n.$$

Additionally, the boundary conditions for h will be specified in $\mathcal{F}$. We consider Dirichlet conditions of the form $h = 0$ almost everywhere on $\partial\Omega$ (in fact, because of the regularity of h, this condition holds everywhere on $\partial\Omega$), and set

$$\mathcal{F} = \mathbf{H}_0^1(\Omega) = \left(W_0^{1,2}(\Omega)\right)^n.$$

Because of the special form of $\mathcal{R}(h)$, the corresponding regularization operator is a second order differential one, and we therefore will need the space

$$\mathbf{H}^2(\Omega) = \left(W^{2,2}(\Omega)\right)^n$$

for the definition of its domain.

3.2 Linearized elasticity

The family that we consider is inspired from the equilibrium equations of linearized elasticity (we refer to [9] for a formal study of three-dimensional elasticity) obtained by defining φ in equation 8 by

$$\varphi(Dh) = \frac{1}{2}\left(\xi \mathrm{Tr}(Dh^T Dh) + (1-\xi)\mathrm{Tr}(Dh)^2\right), \tag{9}$$

where $1/2 < \xi \leq 1$. It is straightforward to verify that the Euler-Lagrange equation corresponding to equation 8 in this case is:

$$\nabla_H \mathcal{R}(h) = \mathrm{div}\left(D\varphi(Dh)\right) = \xi\Delta h + (1-\xi)\nabla(\nabla \cdot h)$$

We thus define the corresponding regularization operator as follows.

Definition 3.2.1. *The linear operator* $A : \mathcal{D}(A) \to H$ *is defined as*

$$\begin{cases} \mathcal{D}(A) = \mathbf{H}_0^1(\Omega) \cap \mathbf{H}^2(\Omega), \\ \\ Ah = \nabla_H \mathcal{R}(h) = \xi\Delta h + (1-\xi)\nabla(\nabla \cdot h) \end{cases}$$

for $1/2 < \xi \leq 1$

4 Definition of the data term $\mathcal{J}$

We want to compare the values of the image I_1 in a neighborhood of a voxel x to those of I_2 in the corresponding neighborhood transformed by h. We propose a statistical framework for doing so, in the spirit of block-matching techniques. Local statistics (mean, covariance matrix) has been found to be very useful for warping scalar images, i.e. real-valued images. This idea can be generalized to tensor-valued images as follows.

4.1 Local mean and covariance matrix

Given a voxel x in the volume Ω, the local mean $\hat{\mu}_1(x)$ is defined as one of the minima with respect to its first argument of the following function defined on $S^+ \times \Omega$

$$\mathcal{C}_1(\mu_1, x) = \frac{1}{\mid \Omega \mid} \int_\Omega \mathcal{D}^2(\mu_1, I_1(y)) G_\gamma(x-y)\, dy,$$

where G_γ is a three-dimensional Gaussian with 0-mean and variance γ^2

$$G_\gamma(x) = \frac{1}{(2\pi\gamma^2)^{3/2}} \exp\left(-\frac{|x|^2}{2\gamma^2}\right).$$

$|x|$ is the Euclidean norm of the vector x and $\mathcal{D}$ is the geodesic distance defined in equation 4 between the two elements μ_1 and $I_1(y)$ of S^+ [20]. $\mathcal{C}$ is a weighted average of the squared geodesic distances between the element μ_1 of S^+ and the elements of the image I_1. The amount of locality is controlled by the parameter γ^2, the variance of the Gaussian. We call $\hat{\mu}_1(x)$ the element of S^+ that minimizes $\mathcal{C}_1$. It is the weighted Riemannian mean of the family $I_1(y)$ of elements of S^+, where y varies in Ω, a notion introduced by Grove, Karcher and Ruh [14].

Hence we have

$$\hat{\mu}_1(x) = \operatorname*{argmin}_{\mu_1} \frac{1}{\mid \Omega \mid} \int_\Omega \mathcal{D}^2(\mu_1, I_1(y)) G_\gamma(x-y)\, dy$$

Because of equation 6 we can write an expression for the gradient of the function $\mathcal{C}$ with respect to μ_1, at μ_1:

$$\nabla_{\mu_1} \mathcal{C}_1(\mu_1, x) = \frac{\mu_1}{\mid \Omega \mid} \int_\Omega \log\left(I_1(y)^{-1}\mu_1\right)\, G_\gamma(x-y)\, dy$$

At the minimum $\hat{\mu}_1(x)$, this gradient is equal to 0:

$$\frac{\hat{\mu}_1(x)}{\mid \Omega \mid} \int_\Omega \log\left(I_1(y)^{-1}\hat{\mu}_1(x)\right)\, G_\gamma(x-y)\, dy = 0 \tag{10}$$

An interpretation of this relation is the following. Each of the matrices

$$\beta_1(x,y) \stackrel{def}{=} -\hat{\mu}_1(x) \log\left(I_1(y)^{-1}\hat{\mu}_1(x)\right)\, G_\gamma(x-y)$$

belongs to the tangent space $T_{\hat{\mu}_1(x)}S^+$, a copy of S, the space of symmetric matrices. Since $T_{\hat{\mu}_1(x)}S^+$ is identified to $\mathbb{R}^6$ through the chart φ_3, one can define the covariance matrix of the vectors $\varphi_3(\beta_1(x,y))$, noted $\beta_1(x,y)$ for simplicity, which have zero-mean according to equation 10:

$$\Lambda_1(x) = \frac{1}{\mid \Omega \mid} \int_\Omega \beta_1(x,y)\beta_1^T(x,y)\, dy.$$

This is a twice contravariant tensor defined on $T_{\hat{\mu}_1(x)}S^+$.

Applying the transformation h to the second image I_2, we can define the corresponding quantities. The local mean at the voxel $h(x)$ becomes:

$$\hat{\mu}_2(x,h) = \operatorname*{argmin}_{\mu_2} \mathcal{C}_2(\mu_2, x) = \operatorname*{argmin}_{\mu_2} \frac{1}{\mid \Omega \mid} \int_\Omega \mathcal{D}^2(\mu_2, \mathcal{T}_h(I_2)(y)) G_\gamma(x-y)\, dy,$$

and satisfies

$$\nabla_{\mu_2} \mathcal{C}_2(\hat{\mu}_2(x,h), x) = \frac{\hat{\mu}_2(x,h)}{\mid \Omega \mid} \int_\Omega \log\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right)\, G_\gamma(x-y)\, dy = 0. \tag{11}$$

The tangent vectors to S^+ at $\hat{\mu}_2(x,h)$ are

$$\beta_2(x,y,h) \stackrel{def}{=} -\hat{\mu}_2(x,h) \log\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right)\, G_\gamma(x-y) \tag{12}$$

and their covariance matrix is

$$\Lambda_2(x,h) = \frac{1}{\mid \Omega \mid} \int_\Omega \beta_2(x,y,h) \beta_2^T(x,y,h)\, dy. \tag{13}$$

We now face a difficulty. We would like to compare the tangent vectors $\beta_1(x,y)$ and $\beta_2(x,y,h)$ but this cannot be done in a straightforward manner since they live in two different vector spaces, $T_{\hat{\mu}_1(x)}S^+$ and $T_{\hat{\mu}_2(x,h)}S^+$. In order to compare them, we must either parallel transport the vectors $\beta_1(x,y)$ to $T_{\hat{\mu}_2(x,h)}S^+$ (obtaining the vectors $\tilde{\beta}_1(x,y,h)$) or the vectors $\beta_2(x,y,h)$ to $T_{\hat{\mu}_1(x)}S^+$ (obtaining the vectors $\tilde{\beta}_2(x,y,h)$).

We can then define the covariance matrices $\tilde{\Lambda}_{12}(x,h)$ and $\tilde{\Lambda}_{21}(x,h)$ of the parallel transported vectors $\tilde{\beta}_1(x,y,h)$ and $\tilde{\beta}_2(x,y,h)$, respectively:

$$\tilde{\Lambda}_{12}(x,h) = \frac{1}{\mid \Omega \mid} \int_\Omega \tilde{\beta}_1(x,y,h) \tilde{\beta}_1^T(x,y,h)\, dy \tag{14}$$

and

$$\tilde{\Lambda}_{21}(x,h) = \frac{1}{\mid \Omega \mid} \int_\Omega \tilde{\beta}_2(x,y,h) \tilde{\beta}_2^T(x,y,h)\, dy \tag{15}$$

4.2 Parallel transport

We now detail the process of parallel transport as illustrated on figure 1.

4.2.1 The equations

To parallel transport a vector $\beta_1(x,y)$ from $T_{\hat{\mu}_1(x)}S^+$ along the curve $\mathcal{G}(t)$ such that $\mathcal{G}(0) = \hat{\mu}_1(x)$ and $\mathcal{G}(1) = \hat{\mu}_2(x,h)$ one has to solve the first-order linear differential equation

$$\nabla_{\dot{\mathcal{G}}(t)} \beta(t) = 0 \tag{16}$$

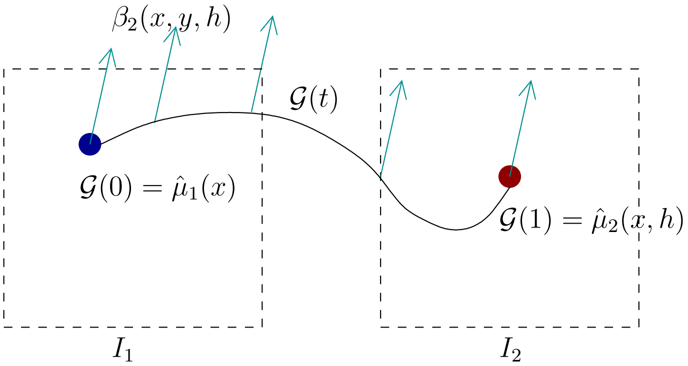


Figure 1: Parallel transport of vector $\beta_2(x,y,h)$ along $\mathcal{G}(t)$

with initial condition $\beta(0)=\beta_1(x,y)$. $\nabla_{\dot{\mathcal{G}}(t)}$ stands for the covariant derivative in the direction $\dot{\mathcal{G}}(t)$ of $T_{\mathcal{G}(t)}S^+$ and equation 16 can be rewritten in local coordinates as:

$$\left(\nabla_{\dot{\mathcal{G}}(t)}\beta(t)\right)^I=\frac{d\,\beta^I}{dt}+\Gamma^I_{JK}(\mathcal{G}(t))\dot{\mathcal{G}}(t)^J\beta^K(t)=0 \tag{17}$$

where the Γ^I_{JK}'s are the Christoffel symbols of the second kind associated to the metric of S^+. This linear differential equation can be written in the form

$$\frac{d\,\beta(t)}{dt}=-\mathcal{A}(t)\beta(t) \tag{18}$$

with the same initial condition $\beta(0)=\beta_1(x,y)$. The 6×6 matrix $\mathcal{A}(t)$ is equal to

$$(\mathcal{A})^I_K(t)=\Gamma^I_{JK}(\mathcal{G}(t))\dot{\mathcal{G}}(t)^J. \tag{19}$$

We recall that $\mathcal{G}(t)$ is the geodesic between $\hat{\mu}_1(x)$ and $\hat{\mu}_2(x,h)$ whose equation is [20]

$$\mathcal{G}(t)=\hat{\mu}_1(x)^{1/2}\exp\left(tX\right)\hat{\mu}_1(x)^{1/2}, \tag{20}$$

where

$$X=\log\left(\hat{\mu}_1(x)^{-1/2}\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1/2}\right) \tag{21}$$

Similar considerations apply to the problem of the parallel transport of the vector $\beta_2(x,y,h)$ along the geodesic $\mathcal{G}(t)$ from $T_{\hat{\mu}_2(x,h)}S^+$to $T_{\hat{\mu}_1(x)}S^+$ by introducing the matrix $\mathcal{B}(t)$.

4.2.2 The matrices $\mathcal{A}$ and $\mathcal{B}$

In the following, we prove that the matrices $\mathcal{A}(t)$ and $\mathcal{B}(t)$ do not depend on the curve parameter t. The solution of equation 18 is therefore

$$\tilde{\beta}_1(x,y,h)=\beta(1)=\exp\left(-\mathcal{A}\right)\beta(0)=\exp\left(-\mathcal{A}\right)\beta_1(x,y) \tag{22}$$

Similarly

$$\tilde{\beta}_2(x,y,h)=\exp\left(-\mathcal{B}\right)\beta_2(x,y,h) \tag{23}$$

We have the following

Proposition 4.2.1. *The matrix $\mathcal{A}(t)$ is independent of t. It is given by the following expression*

$$\mathcal{A}(x,h) = \tag{24}$$
$$-\begin{bmatrix} \psi_1^1 & \psi_1^2/2 & 0 & \psi_1^3/2 & 0 & 0 \\ \psi_2^1 & (\psi_1^1+\psi_2^2)/2 & \psi_1^2 & \psi_2^3/2 & \psi_1^3/2 & 0 \\ 0 & \psi_2^1/2 & \psi_2^2 & 0 & \psi_2^3/2 & 0 \\ \psi_3^1 & \psi_3^2/2 & 0 & (\psi_1^1+\psi_3^3)/2 & \psi_1^2/2 & \psi_1^3 \\ 0 & \psi_3^1/2 & \psi_3^2 & \psi_2^1/2 & (\psi_2^2+\psi_3^3)/2 & \psi_2^3 \\ 0 & 0 & 0 & \psi_3^1/2 & \psi_3^2/2 & \psi_3^3 \end{bmatrix} \stackrel{Def}{=} \mathcal{M}(\psi),$$

where the matrix ψ is equal to $\log\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}\right)$.

The matrix $\mathcal{B}(t)$ is also independent of t and its expression is similar to that of $\mathcal{A}$ by replacing the matrix ψ with the matrix $\theta = \log\left(\hat{\mu}_1(x)^{-1}\hat{\mu}_2(x,h)\right)$.

Proof. It can be shown [32] that the Christoffel symbols, at $\mathcal{G}(t) \in S^+$, are given by:

$$\Gamma\left(E_{pq}, E_{rs}; E_{uv}^*\right) = -\frac{1}{2}\mathrm{tr}\left(E_{pq}\mathcal{G}(t)^{-1}E_{rs}E_{uv}^*\right) - \frac{1}{2}\mathrm{tr}\left(E_{rs}\mathcal{G}(t)^{-1}E_{pq}E_{uv}^*\right)$$
$$\forall p,q,r,s,u,v = 1,2,3$$

The indices p,q,r,s,u,v are used to access the components of the basis elements in matrix form and therefore run from 1 to 3. We introduce the indices I,J,K which run from 1 to 6 since they correspond to the coordinates of a given matrix expressed in the associated local coordinate system (see, for example, equations 17 or 19). Hence we use the following convention:

$$\begin{array}{lll} K = 3(r-1)+s & \text{if} & r \le s \\ J = 3(p-1)+q & \text{if} & p \le q \\ I = 3(u-1)+v & \text{if} & u \le v \end{array} \tag{25}$$

We can now express the quantity $\Gamma_{JK}^I\left(\mathcal{G}(t)\right)\dot{\mathcal{G}}^J(t)$ as:

$$\Gamma\left(E_J, E_K; E_I^*\right)\dot{\mathcal{G}}(t)^J = -\frac{1}{2}\mathrm{tr}\left(\dot{\mathcal{G}}(t)^J E_J\mathcal{G}(t)^{-1}E_K E_I^*\right) - \frac{1}{2}\mathrm{tr}\left(E_K\mathcal{G}(t)^{-1}\dot{\mathcal{G}}(t)^J E_J E_I^*\right)$$

Noting that $\dot{\mathcal{G}}(t)^J E_J = \dot{\mathcal{G}}(t)$, this reduces to:

$$\Gamma\left(E_J, E_K; E_I^*\right)\dot{\mathcal{G}}(t)^J = -\frac{1}{2}\mathrm{tr}\left(\dot{\mathcal{G}}(t)\mathcal{G}(t)^{-1}E_K E_I^*\right) - \frac{1}{2}\mathrm{tr}\left(E_K\mathcal{G}(t)^{-1}\dot{\mathcal{G}}(t)E_I^*\right)$$

In our case $\mathcal{G}(t) \in S^+$, and since

$$\dot{\mathcal{G}}(t) = \hat{\mu}_1(x)^{1/2}X\exp\left(tX\right)\hat{\mu}_1(x)^{1/2} = \hat{\mu}_1(x)^{1/2}\exp\left(tX\right)X\hat{\mu}_1(x)^{1/2}$$

and

$$\mathcal{G}^{-1}(t) = \hat{\mu}_1(x)^{-1/2} \exp(-tX) \hat{\mu}_1(x)^{-1/2}$$

we have

$$\dot{\mathcal{G}}(t)\mathcal{G}^{-1}(t) \stackrel{def}{=} \psi(x,h) = \hat{\mu}_1(x)^{1/2} X \hat{\mu}_1(x)^{-1/2} = \log\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}\right)$$

$$\mathcal{G}^{-1}(t)\dot{\mathcal{G}}(t) \stackrel{def}{=} \theta(x,h)^T = \hat{\mu}_1(x)^{-1/2} X \hat{\mu}_1(x)^{1/2} = \log\left(\hat{\mu}_1(x)^{-1}\hat{\mu}_2(x,h)\right)$$

which do not depend on t. Note that ψ and θ are once contravariant and once covariant tensors.

The last equality in the previous two equations holds because of equation 7.
The matrix $\mathcal{A}(t)$ is therefore independent of t but depends on x and h, thus we note $\mathcal{A}(x,h)$. The contraction of the Christoffel symbols then yields the once covariant, once contravariant tensor:

$$\begin{aligned}
\Gamma\left(E_J, E_K; E_I^*\right) \dot{\mathcal{G}}(t)^J &= \mathcal{A}_K^I(x,h) \\
&= -\frac{1}{2}\mathrm{tr}\left(\psi(x,h)E_K E_I^*\right) - \frac{1}{2}\mathrm{tr}\left(E_K \psi(x,h)^T E_I^*\right) \\
&= -\frac{1}{2}\mathrm{tr}\left(\left(\psi(x,h)E_K + (\psi(x,h)E_K)^T\right) E_I^*\right)
\end{aligned}$$

This provides an expression for $\mathcal{A}(x,h)$, as a function of $\psi(x,h)$, in terms of $\hat{\mu}_1(x)$ and $\hat{\mu}_2(x,h)$ (We use the notation ψ_q^p to denote the pq^{th} element of $\psi(x,h)$). $\square$

4.3 The data term $\mathcal{J}$

We are now ready to define the data term $\mathcal{J}$ in the error criterion for the registration of two DT images I_1 and I_2. This data term is the combination of two terms. The first enforces that the means $\hat{\mu}_1(x)$ and $\hat{\mu}_2(x,h)$ at two corresponding voxels x and $h(x)$ are sufficiently similar. We define the *local* energy:

$$\mathcal{J}_{\mathrm{Mean}}(x,h) = \mathcal{D}^2(\hat{\mu}_2(x,h), \hat{\mu}_1(x)), \tag{26}$$

where $\mathcal{D}^2$ is the geodesic distance 4 in S^+.

We also want the covariance matrices $\Lambda_1(x)$ and $\tilde{\Lambda}_{21}(x,h)$ (respectively $\Lambda_2(x,h)$ and $\tilde{\Lambda}_{12}(x,h)$) to be as close as possible. We therefore define the *local* energy

$$\mathcal{J}_{\mathrm{AC}}(x,h) = \frac{1}{2}\left(|\Lambda_1(x) - \tilde{\Lambda}_{21}(x,h)|_F^2 + |\Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h)|_F^2\right) \tag{27}$$

where $|\cdot|_F$ denotes the Frobenius norm. A more consistent definition of $\mathcal{J}_{\mathrm{AC}}(x,h)$ can be obtained by using the geodesic distance instead of the Frobenius norm

$$\mathcal{J}_{\mathrm{AC}}(x,h) = \frac{1}{2}\left(\mathcal{D}^2\left(\Lambda_1(x), \tilde{\Lambda}_{21}(x,h)\right) + \mathcal{D}^2\left(\Lambda_2(x,h), \tilde{\Lambda}_{12}(x,h)\right)\right) \tag{28}$$

This has very little incidence on the final form of the gradient of $\mathcal{J}_{\mathrm{AC}}$ but may be the source of numerical problems. Indeed, if the region of interest is homogeneous, the covariance matrices tend to be degenerate and the geodesic distance is not defined anymore. In practice, we check if that case happens and only use the Euclidean distance if this is the case. Otherwise, we use the geodesic distance.

We combine these two local criteria into a local data term

$$\mathcal{J}(x,h) = \mathcal{J}_{\text{Mean}}(x,h) + \alpha_1 \mathcal{J}_{\text{AC}}(x,h), \tag{29}$$

where α_1 is a positive weight. The *global* criterion is obtained by integrating the local one over Ω:

$$\mathcal{J}(h, Dh) = \int_\Omega \mathcal{J}(x,h)\, dx = \mathcal{J}_{\text{Mean}}(h, Dh) + \alpha_1 \mathcal{J}_{\text{AC}}(h, Dh). \tag{30}$$

For the sake of clarity, we usually do not express the dependence in Dh of $\mathcal{J}(x,h)$. However, we have to keep in mind that $\mathcal{J}_{\text{Mean}}(x,h)$ and $\mathcal{J}_{\text{AC}}(x,h)$ do depend on the Jacobian of the vector field h because of equation 1. Hence the equation 30.

5 The gradient of the data term

We show that the gradient of the data term exists in H and can be effectively computed and implemented for numerical experiments. The main ingredient in the proof is to show that $\delta_k \mathcal{J}(h, Dh)$ can be written as $(\mathcal{J}_h, k)_H + (\mathcal{J}_{Dh}, Dk)_H$, where $\mathcal{J}_h$ and $\mathcal{J}_{Dh}$ are complicated but computable functions of H. We spend the next sections to prove the following

Theorem 5.0.1. *For any $k \in \mathcal{F}$ the quantity*

$$\delta_k \mathcal{J}(h, Dh) = \lim_{\varepsilon \to 0} \frac{\mathcal{J}(h + \varepsilon k, Dh + \varepsilon Dk) - \mathcal{J}(h, Dh)}{\varepsilon}$$

exists and is equal to

$$(\mathcal{J}_h, k)_H - (\mathcal{J}_{Dh}, Dk)_H,$$

where the functions $\mathcal{J}_h$ and $\mathcal{J}_{Dh}$ are defined in the proof.

5.1 The first variation of $\mathcal{J}_{\text{Mean}}(h, Dh)$

Because of equation 30 we have

$$\delta_k \mathcal{J}_{\text{Mean}}(h, Dh) = \int_\Omega \delta_k \mathcal{J}_{\text{Mean}}(x,h)\, dx.$$

Because of equations 5, 6 and 26 we have

$$\begin{aligned}\delta_k \mathcal{J}_{\text{Mean}}(x,h) = D\mathcal{J}_{\text{Mean}}(x,h)\varphi_3(\delta_k \hat{\mu}_2(x,h)) = \\ G\left(\varphi_3(\hat{\mu}_2(x,h) \log\left((\hat{\mu}_1(x))^{-1} \hat{\mu}_2(x,h)\right))\right) \varphi_3(\delta_k \hat{\mu}_2(x,h)),\end{aligned} \tag{31}$$

where $D\mathcal{J}_{\text{Mean}}(x,h)$ is the differential with respect to its first argument of the function $\mathcal{D}^2(\hat{\mu}_2(x,h),\hat{\mu}_1(x))$. Note that

$$G\left(\varphi_3(\hat{\mu}_2(x,h)\log\left((\hat{\mu}_1(x))^{-1}\,\hat{\mu}_2(x,h)\right))\right)$$

is a covariant vector of $T_{\hat{\mu}_2(x,h)}S^+$ while $\varphi_3(\delta_k\hat{\mu}_2(x,h))$ is a contravariant vector. We thus need to compute $\delta_k\hat{\mu}_2(x,h)$.

5.1.1 Computation of $\delta_k\,\hat{\mu}_2(x,h)$

We remember that the minimum of the functional

$$\mathcal{C}_2(\mu_2,x)=\frac{1}{\mid\Omega\mid}\int_\Omega \mathcal{D}^2(\mu_2,\mathcal{T}_h(I_2)(y))G_\gamma(x-y)\,dy,$$

is achieved at $\hat{\mu}_2(x,h)$. It therefore satisfies

$$\nabla_{\mu_2}\mathcal{C}_2(\hat{\mu}_2,x)=0.$$

To simplify notation, we note $\mathcal{L}$ the vector $\nabla_{\mu_2}\mathcal{C}_2$ of $T_{\hat{\mu}_2}S^+$. This vector is a 3×3 symmetric matrix which we identify with its image by φ_3, a vector of $\mathbb{R}^6$. The previous equation becomes

$$\mathcal{L}(\hat{\mu}_2(x,h),x,h)=0,$$

where the notation indicates that $\mathcal{L}$ depends upon h indirectly through $\hat{\mu}_2$ and directly through its definition. We compute δ_k of the lefthand side and equal it to zero.

$$\frac{\partial\mathcal{L}}{\partial\mu_2}(\hat{\mu}_2,x,h)\varphi_3(\delta_k\hat{\mu}_2(x,h))+\delta_k\mathcal{L}(\hat{\mu}_2,x,h)=0$$

We next compute $\frac{\partial\mathcal{L}}{\partial\mu_2}$, a once contravariant and once covariant tensor, a 6×6 matrix, as well as $\delta_k\mathcal{L}$. For the sake of clarity we drop in the sequel the index 2 in μ_2.

Computation of $\frac{\partial\mathcal{L}}{\partial\mu}$: According to equation 11 we have

$$\begin{aligned}\varphi_3^{-1}\left(\frac{\partial\mathcal{L}}{\partial\mu^I}\right)=&\frac{E_I}{\mid\Omega\mid}\int_\Omega\log\left(\mathcal{T}_h(I_2)^{-1}(y)\mu\right)\,G_\gamma(x-y)\,dy+\\&\frac{\mu}{\mid\Omega\mid}\int_\Omega \mathrm{dlog}\left(\mathcal{T}_h(I_2)^{-1}(y)\mu,\frac{\partial\left(\mathcal{T}_h(I_2)^{-1}(y)\mu\right)}{\partial\mu^I}\right)\,G_\gamma(x-y)\,dy,\;I=1,\cdots,6,\end{aligned}$$

where the notation E_I has been defined in section 4.2.2. We also have

$$\frac{\partial\mathcal{T}_h(I_2)^{-1}(y)\mu}{\partial\mu^I}=\mathcal{T}_h(I_2)^{-1}(y)E_I.$$

Computation of $\delta_k \mathcal{L}$: $\delta_k \mathcal{L}$ is the sum of two terms corresponding to the variation with respect to h and Dh, respectively. We note them $(\delta_k \mathcal{L})^1$ and $(\delta_k \mathcal{L})^2$. We have

$$\varphi_3^{-1}((\delta_k \mathcal{L})^1) = \frac{\mu}{\mid \Omega \mid} \int_\Omega \mathrm{dlog} \left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \frac{\partial \left(\mathcal{T}_h(I_2)^{-1}(y)\mu\right)}{\partial h^i(y)} k^i(y) \right) G_\gamma(x-y)\, dy.$$

Because of equation 1,

$$\frac{\partial \mathcal{T}_h(I_2)^{-1}(y)\mu}{\partial h^i(y)} = -Dh(y)^T I_2(h(y))^{-1} \left(DI_2(h(y)) \right)_i I_2(h(y))^{-1} Dh(y)\mu,$$

where DI_2 is the twice contravariant once covariant tensor obtained by taking the derivative of I_2 with respect to the space coordinates.

We introduce the thrice covariant tensor $\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}$ such that

$$\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}(y) = -Dh^T(y) I_2(h(y))^{-1} DI_2(h(y)) I_2(h(y))^{-1} Dh(y) \tag{32}$$

Because of the linearity of the function $\mathrm{dlog}\,(,)$ with respect to its second argument we obtain

$$\varphi_3^{-1}((\delta_k \mathcal{L})^1) \quad = \quad -\frac{\mu}{\mid \Omega \mid} \int_\Omega \mathrm{dlog} \left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}(y)\mu \right) k(y)\, G_\gamma(x \ - \ y)\, dy. \tag{33}$$

Note that for each value of the covariant index i, the matrix

$$\mu \mathrm{dlog} \left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h} \right)_i (y)\mu \right)$$

is symmetric. At this point we introduce the corresponding once contravariant and once covariant tensor, noted $\frac{\partial \mathcal{L}}{\partial h}$:

$$\begin{aligned} \left(\frac{\partial \mathcal{L}}{\partial h} \right)_i^{\cdot} (x, y, h) = & \\ & - \frac{G_\gamma(x-y)}{\mid \Omega \mid} \varphi_3 \left(\hat{\mu}_2(x,h) \mathrm{dlog} \left(\mathcal{T}_h(I_2)^{-1}(y) \hat{\mu}_2(x,h), \frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}(y) \hat{\mu}_2(x,h) \right) \right). \end{aligned} \tag{34}$$

We write

$$(\delta_k \mathcal{L})^1(x,h) = \int_\Omega \frac{\partial \mathcal{L}}{\partial h}(x,y,h) k(y)\, dy$$

$(\delta_k \mathcal{L})^2$ is obtained in a similar manner:

$$\varphi_3^{-1}((\delta_k \mathcal{L})^2) = \frac{\mu}{\mid \Omega \mid} \int_\Omega \mathrm{dlog} \left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \frac{\partial \left(\mathcal{T}_h(I_2)^{-1}(y)\mu\right)}{\partial Dh_m^l(y)} Dk_m^l(y) \right) G_\gamma(x-y)\, dy$$

where, because of equation 1,

$$\frac{\partial \mathcal{T}_h(I_2)^{-1}(y)\mu}{\partial Dh^l_m(y)} = 1_{ml} I_2(h(y))^{-1} Dh(y)\mu + Dh(y)^T I_2(h(y))^{-1} 1_{lm}\mu,$$

where $1_{lm}$and 1_{ml} are matrices whose only non zero element is located respectively at line l, row m or line m, row l. We introduce a once contravariant thrice covariant tensor $\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}$ such that

$$\left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}\right)^m_l (y) = 1_{ml} I_2(h(y))^{-1} Dh(y) + Dh(y)^T I_2(h(y))^{-1} 1_{lm} \tag{35}$$

and therefore

$$\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}(y) Dk(y) = \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}\right)^m_l (y) Dk^l_m(y),$$

where l and m vary from 1 to 3. Using again the linearity of the function $\mathrm{dlog}\,(\ ,\ )$ with respect to its second argument we obtain

$$\varphi_3^{-1}((\delta_k\mathcal{L})^2) \quad = \quad \frac{\mu}{\mid \Omega \mid}\int_\Omega \mathrm{dlog}\left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}(y)\mu\right) Dk(y)\, G_\gamma(x\ -\ y)\, dy. \tag{36}$$

Note that for each value of the covariant index l and contravariant index m, the matrix

$$\mu \mathrm{dlog}\left(\mathcal{T}_h(I_2)^{-1}(y)\mu, \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}\right)^m_l (y)\mu\right)$$

is symmetric. At this point we introduce the twice contravariant and once covariant tensor, noted $\frac{\partial \mathcal{L}}{\partial Dh}$, such that

$$\left(\frac{\partial \mathcal{L}}{\partial Dh}\right)^{\cdot\, m}_l (x,y,h) = \\ \frac{G_\gamma(x-y)}{\mid \Omega \mid}\varphi_3\left(\hat{\mu}_2(x,h)\mathrm{dlog}\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h), \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}\right)^m_l (y)\hat{\mu}_2(x,h)\right)\right). \tag{37}$$

We write

$$(\delta_k\mathcal{L})^2(x,h) = \int_\Omega \frac{\partial \mathcal{L}}{\partial Dh}(x,y,h) Dk(y)\, dy$$

This allows us to compute $\delta_k\hat{\mu}_2(x,h)$ if the matrix $\frac{\partial \mathcal{L}}{\partial \mu_2}$ is invertible:

$$\varphi_3(\delta_k\hat{\mu}_2(x,h)) = -\left(\frac{\partial \mathcal{L}}{\partial \mu_2}(x,h)\right)^{-1}\int_\Omega \frac{\partial \mathcal{L}}{\partial h}(x,y,h)k(y)\, dy - \\ \left(\frac{\partial \mathcal{L}}{\partial \mu_2}(x,h)\right)^{-1}\int_\Omega \frac{\partial \mathcal{L}}{\partial Dh}(x,y,h)Dk(y)\, dy \tag{38}$$

all expressions being evaluated at $(\hat{\mu}_2, x, h)$.

We define the once contravariant and once covariant tensor

$$\mathbf{T}(x,y,h) = -\left(\frac{\partial \mathcal{L}}{\partial \mu_2}(x,h)\right)^{-1} \frac{\partial \mathcal{L}}{\partial h}(x,y,h), \tag{39}$$

and the twice contravariant and once covariant tensor

$$\mathbf{U}(x,y,h) = \left(\frac{\partial \mathcal{L}}{\partial \mu_2}(x,h)\right)^{-1} \frac{\partial \mathcal{L}}{\partial Dh}(x,y,h), \tag{40}$$

and rewrite equation 38 in a more compact manner, namely:

$$\varphi_3(\delta_k \hat{\mu}_2(x,h)) = \int_\Omega \mathbf{T}(x,y,h)k(y)\,dy - \int_\Omega \mathbf{U}(x,y,h)Dk(y)\,dy. \tag{41}$$

Using indexes, this is equivalent to

$$(\varphi_3(\delta_k \hat{\mu}_2(x,h)))^I = \int_\Omega \mathbf{T}_l^I(x,y,h)k^l(y)\,dy - \int_\Omega \mathbf{U}_l^{Im}(x,y,h)Dk_m^l(y)\,dy,$$
$$I = 1,\cdots,6,\ l,m = 1,\cdots,3$$

In the next sections we will also need the twice contravariant and once covariant tensor $\varphi_3^{-1}(\mathbf{T})$ and the thrice contravariant and once covariant tensor $\varphi_3^{-1}(\mathbf{U})$ which we note $\mathbf{t}$ and $\mathbf{u}$. The previous formula can be rewritten as

$$(\delta_k \hat{\mu}_2(x,h))^{ij} = -\int_\Omega \mathbf{t}_l^{ij}(x,y,h)k^l(y)\,dy - \int_\Omega \mathbf{u}_l^{ijm}(x,y,h)Dk_m^l(y)\,dy,$$
$$i,j = 1,\cdots,3,\ l,m = 1,\cdots,3 \tag{42}$$

5.1.2 An expression for $\delta_k \mathcal{J}_{\mathrm{MEAN}}(h, Dh)$

We are in a position to prove the following

Proposition 5.1.1. *$\delta_k \mathcal{J}_{\mathrm{MEAN}}(h, Dh)$ exists and is of the form of theorem 5.0.1.*

Proof. We introduce the once covariant tensor

$$\mathbf{t}_{\mathrm{Mean}}(x,y,h) = D\mathcal{J}_{\mathrm{Mean}}(x,h)\mathbf{T}(x,y,h),$$

the once covariant and once contravariant tensor

$$\mathbf{u}_{\mathrm{Mean}}(x,y,h) = D\mathcal{J}_{\mathrm{Mean}}(x,h)\mathbf{U}(x,y,h),$$

and write

$$\delta_k \mathcal{J}_{\mathrm{Mean}}(x,h) = \int_\Omega \mathbf{t}_{\mathrm{Mean}}(x,y,h)k(y)\,dy - \int_\Omega \mathbf{u}_{\mathrm{Mean}}(x)Dk(y)\,dy,$$

or, using indexes

$$\delta_k \mathcal{J}_{\text{Mean}}(x,h) = \int_\Omega \mathbf{t}_{\text{Mean } l}(x,y,h) k^l(y)\, dy - \int_\Omega \mathbf{u}^m_{\text{Mean } l}(x,y,h) Dk^l_m(y)\, dy,$$

where, for example:

$$\mathbf{t}_{\text{Mean } l}(x,y,h) = D\mathcal{J}_{\text{Mean I}}(x,h)\mathbf{T}^I_l(x,y,h)$$

We finally introduce the once covariant tensor

$$\mathbf{T}_{\text{Mean}}(x,h) = \int_\Omega \mathbf{t}_{\text{Mean}}(z,x,h)\, dz,$$

and the once covariant and once contravariant tensor

$$\mathbf{U}_{\text{Mean}}(x,h) = \int_\Omega \mathbf{u}_{\text{Mean}}(z,x,h)\, dz,$$

From this follows the fact that $\delta_k \mathcal{J}_{\text{Mean}}(h, Dh)$ can be written in the form of theorem 5.0.1

$$\delta_k \mathcal{J}_{\text{Mean}}(h, Dh) = \int_\Omega \mathbf{T}_{\text{Mean}}(x,h) k(x)\, dx - \int_\Omega \mathbf{U}_{\text{Mean}}(x,h) Dk(x)\, dx$$

□

It is then possible to rewrite $\delta_k \mathcal{J}_{\text{Mean}}(h, Dh)$ in the form of a scalar product with $k(x)$. Indeed, integrating by part, and using the fact that k is zero on $\partial\Omega$, we obtain

$$(\mathbf{U}_{\text{Mean}}, Dk)_H = -\int_\Omega \text{div}\mathbf{U}_{\text{Mean}}(x,h)k(x)\, dx = -\int_\Omega \frac{\partial}{\partial x^m} \mathbf{U}^m_{\text{Mean } l}(x,h) k^l(x)\, dx$$

with x^m the $m^{th} \in [1,2,3]$ coordinate of the position $x \in \mathbb{R}^3$. Hence,

$$\delta_k \mathcal{J}_{\text{Mean}}(h, Dh) = \int_\Omega \left(\mathbf{T}_{\text{Mean}}(x,h) + \text{div}\mathbf{U}_{\text{Mean}}(x,h)\right) k(x)\, dx$$

and

$$\nabla_H \mathcal{J}_{\text{Mean}}(x,h) = \mathbf{T}_{\text{Mean}}(x,h) + \text{div}\mathbf{U}_{\text{Mean}}(x,h)$$

This is the first contribution to $\nabla_H \mathcal{J}(x,h)$ (see equation 3) to be used in our gradient descent.

5.2 The First Variation of $\mathcal{J}_{\text{AC}}(h)$

We would like to compute the second contribution to $\nabla_H \mathcal{J}(x,h)$, associated to the matching term for covariance matrices. The calculations are similar to the previous ones but very much involved so we detail them in appendix A.

5.3 Conclusion

We are now ready to give the proof of theorem 5.0.1.

Proof. It suffices to combine propositions 5.1.1, A.2.1, A.3.1 and A.4.1 to obtain

$$\delta_k \mathcal{J}(h, Dh) = (\mathbf{T}_{\text{Mean}} + \mathbf{T}^1_{\text{AC}} + \mathbf{T}^2_{\text{AC}} + \mathbf{T}^3_{\text{AC}}, k)_H - (\mathbf{U}_{\text{Mean}} + \mathbf{U}^1_{\text{AC}} + \mathbf{U}^2_{\text{AC}} + \mathbf{U}^3_{\text{AC}}, Dk)_H \quad (43)$$

□

This yields the existence of the gradient of the data term:

Proposition 5.3.1. *The gradient $\nabla_H \mathcal{J}(h)$ of data term $\mathcal{J}(h)$ exists and is given by*

$$\nabla_H \mathcal{J}(h) \quad = \quad \mathbf{T}_{\text{Mean}} \ + \ \mathbf{T}^1_{\text{AC}} \ + \ \mathbf{T}^2_{\text{AC}} \ + \ \mathbf{T}^3_{\text{AC}} \ + \ \text{div}\left(\mathbf{U}_{\text{Mean}} + \mathbf{U}^1_{\text{AC}} + \mathbf{U}^2_{\text{AC}} + \mathbf{U}^3_{\text{AC}}\right)$$

Proof. This is a direct consequence of the proof of theorem 5.0.1 and of equation 43. □

We thus know how to compute the gradient of the data and regularization terms. In the next section, we address some of the many numerical difficulties arising in the implementation of this registration technique. We also provide examples on two dimensional synthetic datasets, as a proof of concept.

6 Numerical experiments

In this section, we will illustrate our method on three different examples. An in-depth study of many numerical aspects of this problem is still needed but we believe that current results, though simple, demonstrate the feasibility of the approach. The code was written in C++.

Up to now, we concentrated on the implementation of the gradient of the matching term $\mathcal{J}_{\text{Mean}}(x, h)$. The gradient of the other term $\mathcal{J}_{\text{AC}}(x, h)$ is much more tricky to compute and, most importantly, extremely time consuming because of the numerous numerical integrations to perform. The examples below were thus generated by only using $\nabla_H \mathcal{J}_{\text{Mean}}(x, h)$, which makes sense since we definitely want the local means to match before the local covariance matrices do. Regarding the computation of the gradient of the linear elasticity regularization term, $\nabla_H \mathcal{R}(h)$ (equation 3.2.1), we refer to [17] where adequate numerical schemes were given.

In order to recover large displacements, we used a multiresolution approach. The original images I_1 and I_2 were subsampled such that every level of the multiscale pyramid had a resolution equal to half of the resolution at the previous level. Details on that point can also be found in [17]. In the following experiments, we used 2 levels in the pyramid in addition to the original images. Subsampling was performed by computing local Riemannian averages of I_1 and I_2. Whenever the algorithm converged at a given level, it is easy to propagate the resulting diffeomorphism to the next level by bilinear interpolation. This serves as the initial value of the evolution for the next level. At the lowest resolution, the vector field h is initialized with the identity.

At each resolution, the evolution 3 requires the definition of a few parameters. First of all, we check for convergence by simply looking at the evolution of the energy 30. Whenever it stops decreasing for many iterations, we stop the gradient descent at the current level and propagate the estimated diffeomorphism to the next scale. The scale parameter γ, used to smooth the images at each resolution was fixed, in all our experiments to a small value, typically between 0.5 and 1 pixel.
Finally, the coefficient κ of the regularization term (see equation 8) and the time step dt can be chosen, as proposed in [12], as follows. In order to control the range of the regularization term, κ is normalized by the maximum value $\kappa_0 = |\nabla_H \mathcal{J}(h)|_\infty$ such that $C = \kappa\kappa_0$. Using C instead of κ makes the algorithm much more stable and we used $\kappa = 10$. The time step dt is adapted at each level of the pyramid so that Cdt is less than a specified value. In our experiments, we set $Cdt = 0.2$.

It also makes sense to favor rigid transformations like translations and rotations at coarse resolutions. Consequently, at the first and second levels of the pyramid, we fit the estimated non-rigid deformation field with the best rigid transformation, expressed as the combination of a rotation and a translation, and this rigid transformation is used in place of the estimated non-rigid one. This is easily achieved by solving, for the 2D case, the following linear system

$$\underbrace{\begin{bmatrix} h(x_1)^1 & \cdots & h(x_N)^1 \\ h(x_1)^2 & \cdots & h(x_N)^2 \\ 1 & \cdots & 1 \end{bmatrix}}_{\tilde{X}} = \underbrace{\begin{bmatrix} R_{11} & R_{12} & T^1 \\ R_{12} & R_{22} & T^2 \\ 0 & 0 & 1 \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} x_1^1 & \cdots & x_N^1 \\ x_1^2 & \cdots & x_N^2 \\ 1 & \cdots & 1 \end{bmatrix}}_{X}$$

where, for an image containing N pixels, $\tilde{X}$ is the $3 \times N$ matrix whose columns contain the components of the estimated diffeomorphism, X is the $3 \times N$ matrix whose columns contain the coordinates of each pixel and M is the 3×3 matrix containing the rotation matrix R and the translation T. $\tilde{X}$ and X being given, we simply have $M = \tilde{X}\left(X^T X\right)^{-1} X^T$. A QR decomposition of the submatrix R of M should be used to replace R by only its rotational component.

In figure 2, we present the very simple example of a translation. It is perfectly recovered. In figure 3, the more complicated example of a rotation is shown. We can see that the rotation is well recovered and, most importantly, that the tensor reorientation is correctly performed. Finally, we considered the case of a non-rigid transformation taking an ellipse onto a circle (figure 4). During the evolution, the algorithm first estimated a translation that maximized the overlap of the ellipse on the circle. Afterwards, non-rigid effects took the advantage in order to fully deform the ellipse into the circle.

7 Conclusion

We have presented a mathematical study of the non-rigid registration problem for diffusion tensor images. We setup a variational formulation taking into account the Riemannian structure of the space of diffusion tensors to derive the matching energy. To our knowledge, this is the first time that the properties of the manifold S^+ are exploited for DTI registration. As shown in this paper, the computations are a bit tedious and the numerical implementation must be done carefully. We demonstrated the feasibility of the approach by successfully applying the algorithm to three different transformations.

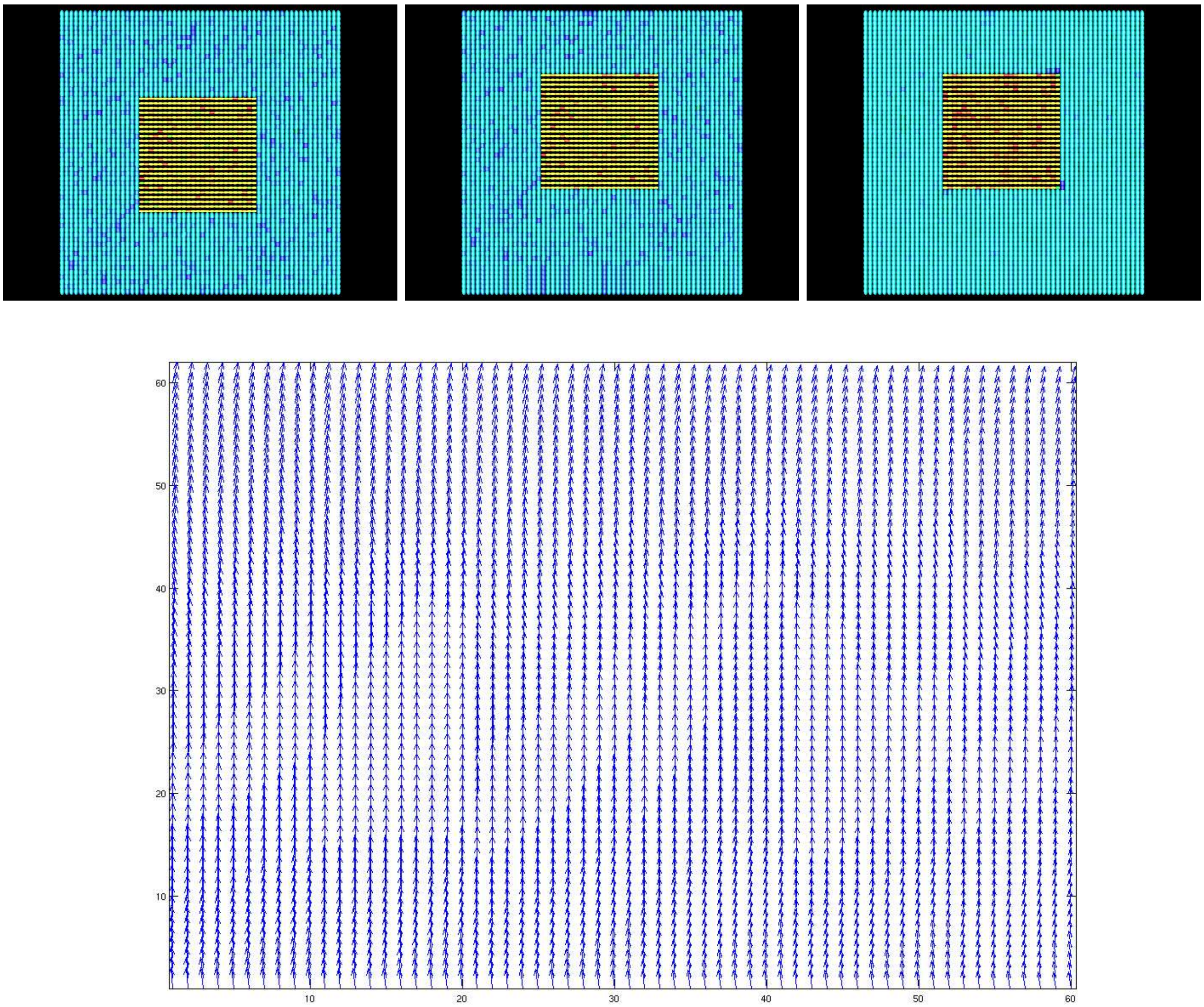
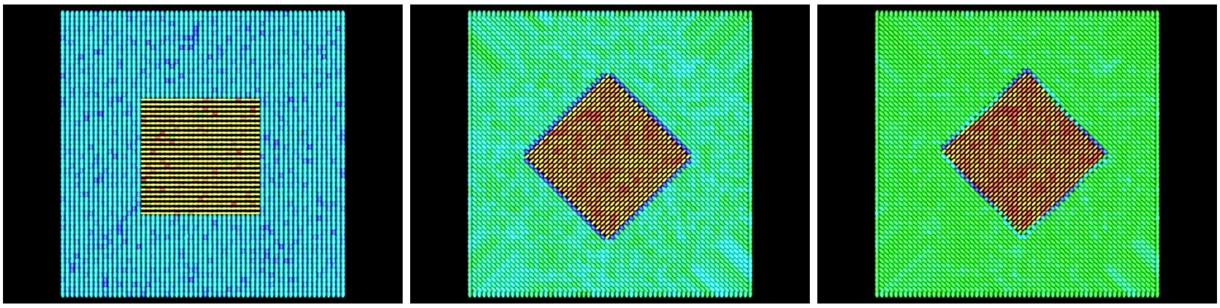


Figure 2: Estimation of the translation (6 pixels) of a square. I_2, I_1 and $\mathcal{T}_h(I_2)$ on top, h at bottom.



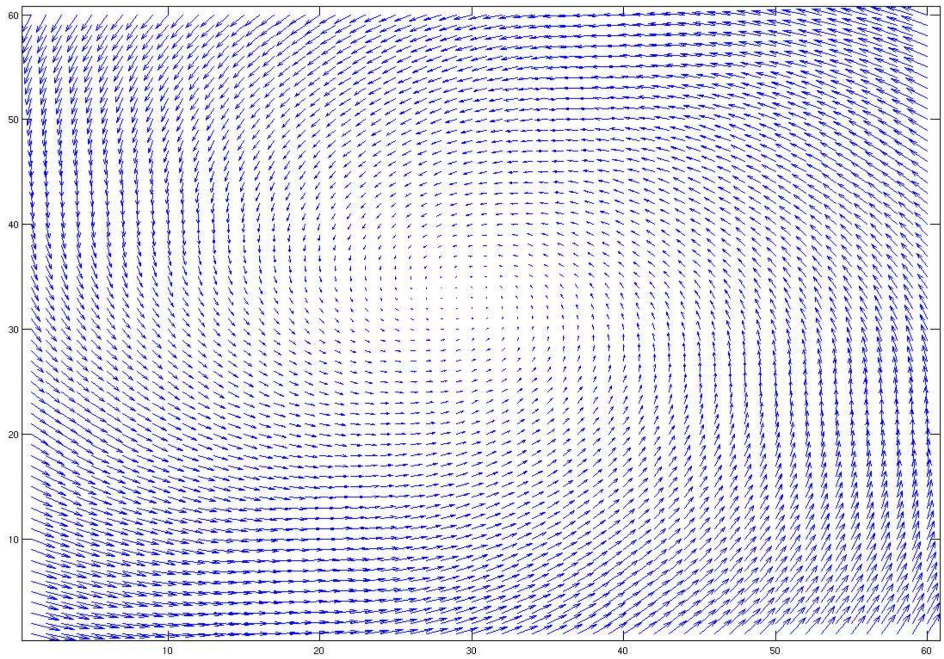


Figure 3: Estimation of the rotation ($\pi/4$) of a square. I_2, I_1 and $\mathcal{T}_h(I_2)$ on top, h at bottom.

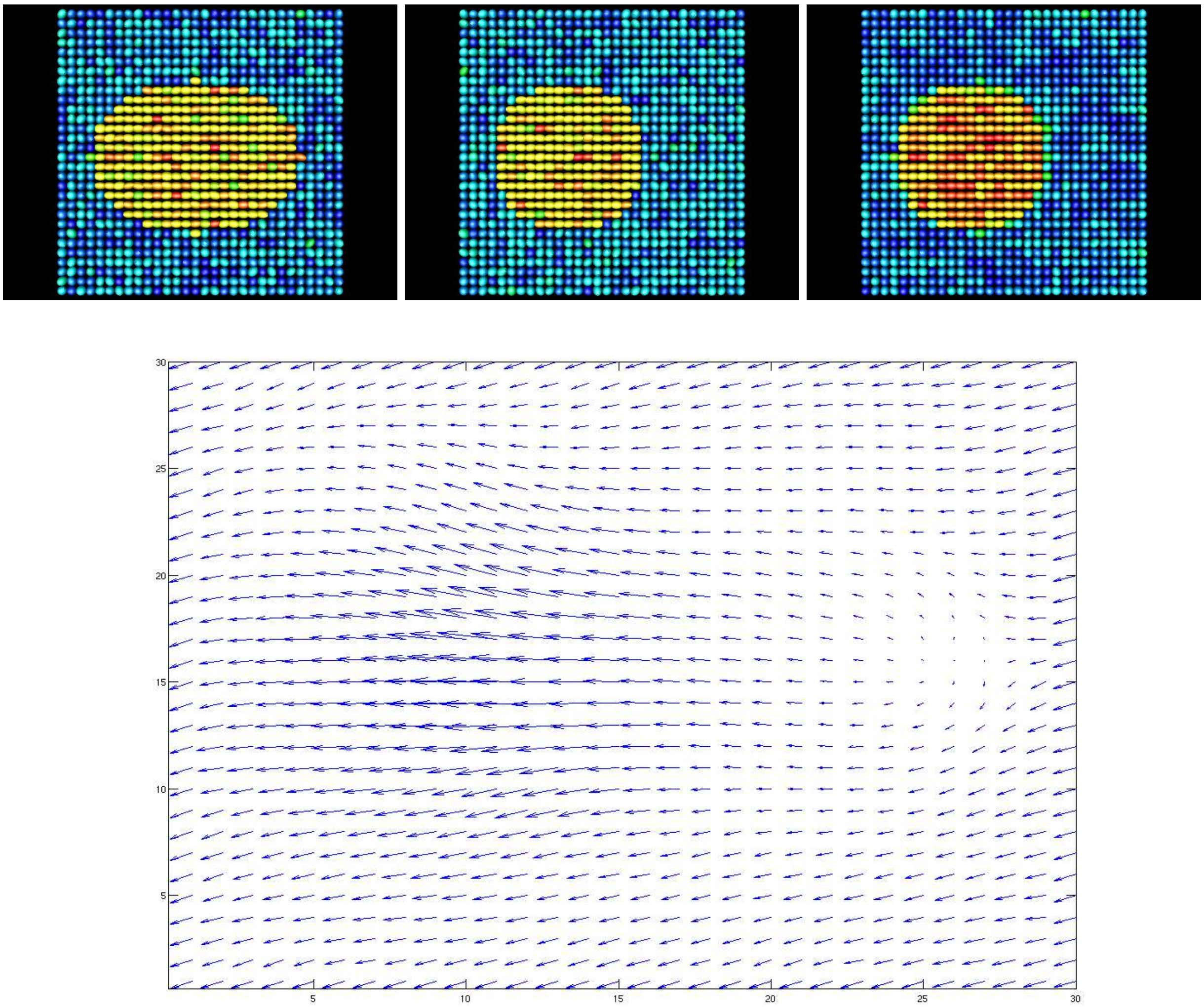


Figure 4: Estimation of the diffeomorphism taking an ellipse onto a circle. I_2, I_1 and $\mathcal{T}_h(I_2)$ on top, h at bottom.

A Details on the first Variation of $\mathcal{J}_{\text{AC}}(h, Dh)$

A.1 Introduction

In this appendix, we compute the first variation of the term $\mathcal{J}_{\text{AC}}(h, Dh)$ introduced in section 4 and corresponding to the matching term, for the DTI registration problem, between local covariance matrices of the images I_1 and $\mathcal{T}_h(I_2)$.
We identify the covariance matrices, elements of $S(6)$ with their images by the canonical map φ_6. Because of (30) we have

$$\delta_k \mathcal{J}_{\text{AC}}(h, Dh) = \int_\Omega \delta_k \mathcal{J}_{\text{AC}}(x, h)\, dx.$$

Because of (27) we have

$$\begin{aligned}\delta_k \mathcal{J}_{\text{AC}}(x,h) = \frac{1}{2}\Big(\quad & \tfrac{\partial}{\partial \Lambda_2} \|\Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h)\|_F^2\, \delta_k \Lambda_2(x,h) + \\ & \tfrac{\partial}{\partial \tilde{\Lambda}_{12}} \|\Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h)\|_F^2\, \delta_k \tilde{\Lambda}_{12}(x,h) \ + \\ & \tfrac{\partial}{\partial \tilde{\Lambda}_{21}} \|\Lambda_1(x) - \tilde{\Lambda}_{21}(x,h)\|_F^2\, \delta_k \tilde{\Lambda}_{21}(x,h) \Big)\end{aligned}$$

In this equation, the partial derivatives are covariant vectors and the variations $\delta_k \cdot$ are contravariant vectors. The expression of the partial derivatives follows from the fact that $\|A - B\|_F^2 = \operatorname{tr}\big((A-B)(A-B)^T\big)$ and $\frac{\partial}{\partial A}\|A-B\|_F^2 = A - B = -\frac{\partial}{\partial B}\|A-B\|_F^2$

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial \Lambda_2}\|\Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h)\|_F^2 &= \Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h) \stackrel{def}{=} \Theta(x) \\ \frac{1}{2}\frac{\partial}{\partial \tilde{\Lambda}_{12}}\|\Lambda_2(x,h) - \tilde{\Lambda}_{12}(x,h)\|_F^2 &= \tilde{\Lambda}_{12}(x,h) - \Lambda_2(x,h) = -\Theta(x) \\ \frac{1}{2}\frac{\partial}{\partial \tilde{\Lambda}_{21}}\|\Lambda_1(x) - \tilde{\Lambda}_{21}(x,h)\|_F^2 &= \tilde{\Lambda}_{21}(x,h) - \Lambda_1(x) \stackrel{def}{=} \Phi(x)\end{aligned}$$

Note that in these formulas, Θ and Φ are 21-dimensional covariant vectors that we identify for convenience with their images by φ_9^{-1}. Θ and Φ are therefore twice covariant tensors. Note that, since we know how to compute the gradient of the geodesic distance function $\mathcal{D}$, it is straightforward to define these quantities when they involve $\mathcal{D}$ instead of the Frobenius norm.

We have obtained an expression for $\delta_k \mathcal{J}_{\text{AC}}(x,h)$ and thus, by integration, $\delta_k \mathcal{J}_{\text{AC}}(h, Dh)$ as the sum of three terms

$$\begin{aligned}\delta_k \mathcal{J}^1_{\text{AC}}(x,h) &= \Theta(x)\delta_k \Lambda_2(x,h) \\ \delta_k \mathcal{J}^2_{\text{AC}}(x,h) &= -\Theta(x)\delta_k \tilde{\Lambda}_{12}(x,h) \\ \delta_k \mathcal{J}^3_{\text{AC}}(x,h) &= \Phi(x)\delta_k \tilde{\Lambda}_{21}(x,h)\end{aligned}$$

Let us write the first equation using indexes:

$$\delta_k \mathcal{J}^1_{\text{AC}}(x,h) = \Theta_{IJ}\delta_k \Lambda_2^{IJ}(x,h)$$

We thus need to compute the three quantities $\delta_k\Lambda_2(x,h)$, $\delta_k\tilde{\Lambda}_{12}(x,h)$, $\delta_k\tilde{\Lambda}_{21}(x,h)$. It involves taking derivatives of logarithms and exponentials of matrices which require some numerical care. In [21], we detailed how to evaluate the directional derivative of matrix logarithms and we already referred the reader to [24] for details on the exponential case. The following computations are not difficult but tend to be a little bit involved.

A.2 Computation of $\delta_k\Lambda_2(x,h)$

Because of (13) we have

$$\delta_k\Lambda_2(x,h) = \frac{1}{\mid\Omega\mid}\int_\Omega \Big((\delta_k\beta_2(x,y,h))\beta_2^T(x,y,h) + \beta_2(x,y,h)(\delta_k\beta_2(x,y,h))^T\Big)\,dy \tag{44}$$

Because of (12) we have

$$\begin{aligned}\delta_k\beta_2(x,y,h) = -G_\gamma(x-y)\varphi_3\Big(\,&(\delta_k\hat{\mu}_2(x,h))\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\big) + \\ &\hat{\mu}_2(x,h)\left(\delta_k\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\big)\right)\Big)\end{aligned} \tag{45}$$

We have computed $\delta_k\hat{\mu}_2(x,h)$ in section 5.1.1, we now compute $\delta_k\log(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h))$.

Computation of $\delta_k\log(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h))$**:** We note that since the matrix $\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)$ is similar to $\hat{\mu}_2(x,h)^{1/2}\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)^{1/2}$ which belongs to S^+, it satisfies the hypotheses of corollary 2.2.2 and we can write

$$\delta_k\left(\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\big)\right) \quad = \quad \mathrm{dlog}\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h), \delta_k\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right)\right)$$

We need to compute $\delta_k(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h))$. Using the formula for the derivative of a product

$$\delta_k\left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right) = \delta_k\left(\mathcal{T}_h(I_2)^{-1}(y)\right)\hat{\mu}_2(x,h) + \mathcal{T}_h(I_2)^{-1}(y)\delta_k\hat{\mu}_2(x,h)$$

We have already computed the second term in the righthand side (equations (32) and (35)), hence

$$\delta_k\left(\mathcal{T}_h(I_2)^{-1}(y)\right)\hat{\mu}_2(x,h) = \frac{\partial\mathcal{T}_h(I_2)^{-1}}{\partial h}(y)\hat{\mu}_2(x,h)k(y) + \frac{\partial\mathcal{T}_h(I_2)^{-1}}{\partial Dh}(y)\hat{\mu}_2(x,h)Dk(y)$$

Hence we get,

$$\begin{aligned}\delta_k\ \left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right) &= \mathcal{T}_h(I_2)^{-1}(y)\delta_k\,\hat{\mu}_2(x,h) \\ &\quad + \frac{\partial\mathcal{T}_h(I_2)^{-1}}{\partial h}(y)\hat{\mu}_2(x,h)k(y) + \frac{\partial\mathcal{T}_h(I_2)^{-1}}{\partial Dh}(y)\hat{\mu}_2(x,h)Dk(y),\end{aligned}$$

which we write in tensor form

$$\begin{aligned}&\delta_k\ \left(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\right) = \\ &\quad\int_\Omega \mathbf{t}_1(x,y,z)k(z)\,dz + \mathbf{t}_2(x,y)k(y) - \int_\Omega \mathbf{u}_1(x,y,z)Dk(z)\,dz - \mathbf{u}_2(x,y)Dk(y),\end{aligned} \tag{46}$$

where $\mathbf{t}_1$ is the once contravariant and twice covariant tensor obtained by contracting the second covariant index of $\mathcal{T}_h(I_2)^{-1}(y)$ with the first contravariant index of $\mathbf{t}$:

$$\mathbf{t}_1{}^{i}_{jl}(x,y,z) = -(\mathcal{T}_h(I_2)^{-1})_{jm}(y)\mathbf{t}_l^{mi}(x,z),$$

and $\mathbf{u}_1$ is the twice covariant and twice contravariant tensor obtained from $\mathbf{U}$ in a similar fashion:

$$\mathbf{u}_1{}^{im}_{jl}(x,y,z) = -(\mathcal{T}_h(I_2)^{-1})_{jn}(y)\mathbf{u}_l^{nim}(x,z).$$

The tensor $\mathbf{t}_2 = \frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}(y)\hat{\mu}_2$ is once contravariant and twice covariant; its coordinates are given by

$$\mathbf{t}_2{}^{i}_{jl} = \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial h}\right)_{jlm} \hat{\mu}_2^{mi}.$$

The tensor $\mathbf{u}_2 = \frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}(y)\hat{\mu}_2$ is twice contravariant and twice covariant; its coordinates are given by

$$\mathbf{u}_2{}^{im}_{jl} = \left(\frac{\partial \mathcal{T}_h(I_2)^{-1}}{\partial Dh}\right)^m_{jln} \hat{\mu}_2^{ni}.$$

Computation of $\delta_k\beta_2$ and $\delta_k\Lambda_2$: We now do a bit of rewriting in order to get an expression for $\delta_k\beta_2$ and $\delta_k\Lambda_2$. This is tedious but not difficult. We first prove the following

Lemma A.2.1. *$\delta_k\beta_2$ can be written as*

$$\delta_k\beta_2(x,y) = \int_\Omega \mathbf{T}_{1\beta_2}(x,y,z)k(z)\,dz + \mathbf{T}_{2\beta_2}(x,y)k(y) - \int_\Omega \mathbf{U}_{1\beta_2}(x,y,z)Dk(z)\,dz - \mathbf{U}_{2\beta_2}(x,y)Dk(y),$$

where the expressions of the tensors $\mathbf{T}_{1\beta_2}$, $\mathbf{T}_{2\beta_2}$, $\mathbf{U}_{1\beta_2}$ and $\mathbf{U}_{2\beta_2}$ are given in the proof.

Proof. We combine equations (41) and (46). Using equation (45) we can then write

$$\begin{aligned}\delta_k\beta_2(x,y) = &\\ -\varphi_3(\Big(\Big(\int_\Omega \mathbf{t}(x,z)k(z)\,dz - \int_\Omega \mathbf{u}(x,z)Dk(z)\,dz\Big)\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\big) + &\\ \hat{\mu}_2(x,h)\Big(\int_\Omega \mathbf{t}_1(x,y,z)k(z)\,dz + \mathbf{t}_2(x,y)k(y) - &\\ \int_\Omega \mathbf{u}_1(x,y,z)Dk(z) - \mathbf{u}_2(x,y)Dk(y)\Big)\Big)G(x-y)),\end{aligned}$$

and obtain

$$\begin{aligned}\mathbf{T}_{1\beta_2}(x,y,z) &= -\varphi_3((\mathbf{t}(x,z)\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x)\big) + \hat{\mu}_2(x,h)\mathbf{t}_1(x,y))G(x-y))\\ \mathbf{T}_{2\beta_2}(x,y) &= -\varphi_3(\hat{\mu}_2(x,h)\mathbf{t}_2(x,y)G(x-y))\\ \mathbf{U}_{1\beta_2}(x,y,z) &= \varphi_3((\mathbf{u}(x,z)\log\big(\mathcal{T}_h(I_2)^{-1}(y)\hat{\mu}_2(x,h)\big) + \hat{\mu}_2(x,h)\mathbf{u}_1(x,y))G(x-y))\\ \mathbf{U}_{2\beta_2}(x,y) &= \varphi_3(\hat{\mu}_2(x,h)\mathbf{u}_2(x,y)G(x-y))\end{aligned}$$

□

This lemma allows us to prove the following proposition concerning the form of the first term in the expression of $\delta_k \mathcal{J}_{AC}$.

Proposition A.2.1. *The first term $\delta_k \mathcal{J}^1_{\mathrm{AC}}$ in the expression of $\delta_k \mathcal{J}_{\mathrm{AC}}$ is of the form described in theorem 5.0.1.*

Proof. The previous manipulations and equation (44) yield

$$
\begin{aligned}
&(\delta_k \Lambda_2(x))^{IJ} = \\
&\quad \frac{1}{\mid \Omega \mid} \int_\Omega \Big(\Big(\int_\Omega \mathbf{T}^I_{1\beta_2\ l}(x,y,z) k^l(z)\, dz + \mathbf{T}^I_{2\beta_2\ l}(x,y) k^l(y) \\
&\quad - \int_\Omega \mathbf{U}^{Im}_{2\beta_2\ l}(x,y,z) Dk^l_m(z)\, dz - \mathbf{U}^{Im}_{1\beta_2\ l}(x,y) Dk^l_m(y) \Big) \beta^J_2(x,y) \\
&\quad + \beta^I_2(x,y) \Big(\int_\Omega \mathbf{T}^J_{1\beta_2\ l}(x,y,z) k^l(z)\, dz + \mathbf{T}^J_{2\beta_2\ l}(x,y) k^l(y) \\
&\quad - \int_\Omega \mathbf{U}^{Jm}_{1\beta_2\ l}(x,y,z) Dk^l_m(z)\, dz - \mathbf{U}^{Jm}_{2\beta_2\ l}(x,y) Dk^l_m(y) \Big) \Big)\, dy
\end{aligned}
$$

The corresponding term $\delta_k \mathcal{J}^1_{\mathrm{AC}}(x,h) = \Theta_{IJ}(x) (\delta_k \Lambda_2(x))^{IJ}$ in $\delta_k \mathcal{J}_{\mathrm{AC}}(x,h)$ is

$$
\begin{aligned}
&\frac{\Theta_{IJ}(x)}{\mid \Omega \mid} \int_\Omega \Big(\Big(\int_\Omega \mathbf{T}^I_{1\beta_2\ l}(x,y,z) k^l(z)\, dz + \mathbf{T}^I_{2\beta_2\ l}(x,y) k^l(y) \\
&\quad - \int_\Omega \mathbf{U}^{Im}_{1\beta_2\ l}(x,y,z) Dk^l_m(z)\, dz - \mathbf{U}^{Im}_{l}(x,y) Dk^{2\beta_2\ l}_m(y) \Big) \beta^J_2(x,y) \\
&\quad + \beta^I_2(x,y) \Big(\int_\Omega \mathbf{T}^J_{1\beta_2\ l}(x,y,z) k^l(z)\, dz + \mathbf{T}^J_{2\beta_2\ l}(x,y) k^l(y) \\
&\quad - \int_\Omega \mathbf{U}^{Jm}_{1\beta_2\ l}(x,y,z) Dk^l_m(z)\, dz - \mathbf{U}^{Jm}_{2\beta_2\ l}(x,y) Dk^l_m(y) \Big) \Big)\, dy
\end{aligned} \tag{47}
$$

We define

$$
\begin{aligned}
\mathbf{H}_l(x,y) = \frac{\Theta_{IJ}(x)}{\mid \Omega \mid} \int_\Omega \Big(\mathbf{T}^I_{1\beta_2\ l}(x,z,y) \beta^J_2(x,z) + \beta^I_2(x,z) \mathbf{T}^J_{1\beta_2\ l}(x,z,y) \Big)\, dz + \\
\frac{\Theta_{IJ}(x)}{2 \mid \Omega \mid} \Big(\mathbf{T}^I_{2\beta_2\ l}(x,y) \beta^J_2(x,y) + \beta^I_2(x,y) \mathbf{T}^J_{2\beta_2\ l}(x,y) \Big),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{K}^m_l(x,y) = \frac{\Theta_{IJ}(x)}{\mid \Omega \mid} \int_\Omega \Big(\mathbf{U}^{Im}_{1\beta_2\ l}(x,z,y) \beta^J_2(x,z) + \beta^I_2(x,z) \mathbf{U}^{Jm}_{1\beta_2\ l}(x,z,y) \Big)\, dz + \\
\frac{\Theta_{IJ}(x)}{2 \mid \Omega \mid} \Big(\mathbf{U}^{Im}_{2\beta_2\ l}(x,y) \beta^J_2(x,y) + \beta^I_2(x,y) \mathbf{U}^{Jm}_{2\beta_2\ l}(x,y) \Big),
\end{aligned}
$$

and rewrite (47) as

$$\delta_k \mathcal{J}^1_{\mathrm{AC}}(x,h) = \int_\Omega \mathbf{H}_l(x,y,h)k^l(y)\,dy - \int_\Omega \mathbf{K}_l^m(x,y,h)Dk^l_m(y)\,dy$$

The corresponding term in $\delta_k \mathcal{J}_{\mathrm{AC}}(h)$ is

$$\int_\Omega \left(\int_\Omega \mathbf{H}_l(x,y,h)k^l(y)\,dy \right)\,dx - \int_\Omega \left(\int_\Omega \mathbf{K}_l^m(x,y,h)Dk^l_m(y)\,dy \right)\,dx. \tag{48}$$

Now define

$$\overline{\mathbf{H}}_l(x,h) = \int_\Omega \mathbf{H}_l(z,x,h)\,dz,$$

and

$$\overline{\mathbf{K}}_l^m(x,h) = \int_\Omega \mathbf{K}_l^m(z,x,h)\,dz.$$

Exchanging the order of summation in the second term of (48) and renaming the variables, we obtain a new form of (48):

$$\int_\Omega \overline{\mathbf{H}}_l(x,h)k^l(x)\,dx - \int_\Omega \overline{\mathbf{K}}_l^m(x,h)Dk^l_m(x)\,dx \overset{Def}{=} \int_\Omega \mathbf{T}^1_{\mathrm{AC}}(x,h)k(x)\,dx - \int_\Omega \mathbf{U}^1_{\mathrm{AC}}(x,h)Dk(x)\,dx$$

□

A.3 Computation of $\delta_k \tilde{\Lambda}_{12}(x,h)$

According to equation (14), we have

$$\delta_k \tilde{\Lambda}_{12}(x,h) = \frac{1}{\mid \Omega \mid} \int_\Omega \left(\delta_k \tilde{\beta}_1(x,y,h) \tilde{\beta}_1^T(x,y,h) + \tilde{\beta}_1(x,y,h) \left(\delta_k \tilde{\beta}_1(x,y,h) \right)^T \right)\,dy$$

Because of (22), we have

$$\delta_k \tilde{\beta}_1(x,y,h) = \delta_k \Big(\exp\left(-\mathcal{A}(x,h)\right) \beta_1(x,y) \Big) = \Big(\delta_k \exp\left(-\mathcal{A}(x,h)\right) \Big) \beta_1(x,y), \tag{49}$$

since the vectors $\beta_1(x,y)$ are not functions of h. $\mathcal{A}(x,h)$ is defined in section 4.2. In order to compute $\delta_k \exp\left(-\mathcal{A}(x,h)\right)$, we use the following result from [24]. Let X be a diagonalizable matrix of $M_n(\mathbb{R})$, and V a matrix. We are interested in computing the directional derivative of the exponential of X in the direction V

$$\mathrm{dexp}\,(X,V) = \lim_{t\to 0} \frac{1}{t} (\exp(X+tV) - \exp(X))$$

The following theorem (page 41 of [24]) provides an answer and shows that $\mathrm{dexp}\,(X,V)$ is, like the matrix logarithm, linear in its second argument V.

Theorem A.3.1. *If $X = ZDZ^{-1}$ is the spectral decomposition of the semi-simple matrix X, its directional derivative in the direction V is given by*

$$\mathrm{dexp}\,(X, V) = Z\left(\overline{V} \bullet \Xi\right) Z^{-1}$$

where $\overline{V} = Z^{-1}VZ$ and $\overline{V} \bullet \Xi$ denote the Hadamard (entry-by-entry) product of $\overline{V}$ with the matrix Ξ whose entries are given by:

$$\Xi^i_j = \Xi^j_i = \begin{cases} \frac{e^{\lambda_i} - e^{\lambda_j}}{\lambda_i - \lambda_j} & if \lambda_i \neq \lambda_j \\ e^{\lambda_i} & if \lambda_i = \lambda_j \end{cases}$$

Computation of $\delta_k \exp\left(-\mathcal{A}(x,h)\right)$: According to definition 2.2.1 and the chain rule

$$\delta_k \exp\left(-\mathcal{A}(x,h)\right) = -\mathrm{dexp}\left(-\mathcal{A}(x,h), \delta_k \mathcal{A}(x,h)\right).$$

According to equation (24), in order to compute $\delta_k \mathcal{A}(x,h)$, we need to compute $\delta_k \psi(x,h)$ where

$$\psi(x,h) = \log\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}\right).$$

According to corollary 2.2.2 and definition 2.2.1 we can write

$$\delta_k \psi(x,h) = \mathrm{dlog}\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}, \left(\delta_k \hat{\mu}_2(x,h)\right)\hat{\mu}_1(x)^{-1}\right).$$

Because $\mathrm{dlog}\,(,)$ is a linear function of its second argument, using equation (41), the previous equation can be rewritten as

$$\delta_k \psi(x,h) = \int_\Omega \mathbf{t}_\psi(x,z)k(z)\,dz - \int_\Omega \mathbf{u}_\psi(x,z)Dk(z)\,dz,$$

where

$$\begin{aligned} \mathbf{t}_\psi(x,z) &= \mathrm{dlog}\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}, \mathbf{t}(x,z)\hat{\mu}_1(x)^{-1}\right) \\ \mathbf{u}_\psi(x,z) &= \mathrm{dlog}\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}, \mathbf{u}(x,z)\hat{\mu}_1(x)^{-1}\right). \end{aligned}$$

Using indexes,

$$\mathbf{t}_{\psi\;\,l}^{\;\cdot\,\cdot} = \sum_{m=1}^{3} \mathrm{dlog}\left(\hat{\mu}_2(x,h)\hat{\mu}_1(x)^{-1}, \mathbf{T}_l^{\cdot\,m}\left(\hat{\mu}_1(x)^{-1}\right)_{m\,\cdot}\right)$$

Since the relation between ψ and $\mathcal{A}$ is linear ($\mathcal{A} = \mathcal{M}(\psi)$, see equation (24)), we have

$$\delta_k \mathcal{A}(x,h) = \int_\Omega \mathcal{M}(\mathbf{t}_\psi(x,z))k(z)\,dz - \int_\Omega \mathcal{M}(\mathbf{u}_\psi(x,z))Dk(z)\,dz,$$

and therefore, using the linearity in the second argument of $\mathrm{dexp}\,(\cdot,\cdot)$,

$$\begin{aligned} &\delta_k \exp\left(-\mathcal{A}(x,h)\right) = \\ &-\int_\Omega \mathrm{dexp}\left(-\mathcal{A}(x,h), \mathcal{M}(\mathbf{t}_\psi(x,z))\right)k(z)\,dz + \int_\Omega \mathrm{dexp}\left(-\mathcal{A}(x,h), \mathcal{M}(\mathbf{u}_\psi(x,z))\right)Dk(z)\,dz. \end{aligned}$$

We obtain an expression for $\delta_k\tilde{\beta}_1$

$$\delta_k\tilde{\beta}_1(x,y) = \int_\Omega \mathbf{T}_{\tilde{\beta}_1}(x,y,z)k(z)\,dz - \int_\Omega \mathbf{U}_{\tilde{\beta}_1}(x,y,z)Dk(z)\,dz,$$

where

$$\begin{aligned}\mathbf{T}_{\tilde{\beta}_1}(x,y,z) &= -\mathrm{dexp}\left(-\mathcal{A}(x), \mathcal{M}(\mathbf{t}_\psi(x,z))\right)\beta_1(x,y)\\ \mathbf{U}_{\tilde{\beta}_1}(x,y,z) &= -\mathrm{dexp}\left(-\mathcal{A}(x), \mathcal{M}(\mathbf{u}_\psi(x,z))\right)\beta_1(x,y)\end{aligned}$$

This allows us to prove the following

Proposition A.3.1. *The second term, $\delta_k\mathcal{J}^2_{AC}$, in the expression of $\delta_k\mathcal{J}_{AC}$ is of the form described in theorem 5.0.1.*

Proof. The previous manipulations yield

$$\begin{aligned}\Big(\delta_k\tilde{\Lambda}_{12}(x)\Big)^{IJ} = \frac{1}{\mid\Omega\mid}\int_\Omega \bigg(\Big(\int_\Omega \Big(\mathbf{T}^I_{\tilde{\beta}_1\ l}(x,y,z)k^l(z) - \mathbf{U}^{Im}_{\tilde{\beta}_1\ l}(x,y,z)Dk^m_l(z)\Big)\ dz\Big)\tilde{\beta}^J_1(x,y) + \\ \tilde{\beta}^I_1(x,y)\int_\Omega \Big(\mathbf{T}^J_{\tilde{\beta}_1\ l}(x,y,z)k^l(z) - \mathbf{U}^{Jm}_{\tilde{\beta}_1\ l}(x,y,z)Dk^m_l(z)\Big)\ dz\bigg)\,dy\end{aligned}$$

The corresponding term, $\delta_k\mathcal{J}^2_{\mathrm{AC}}(x,h)$, in $\delta_k\mathcal{J}_{\mathrm{AC}}(x,h)$ is

$$\begin{aligned}-\frac{\Theta_{IJ}(x)}{\mid\Omega\mid}\int_\Omega \bigg(\Big(\int_\Omega \Big(\mathbf{T}^I_{\tilde{\beta}_1\ l}(x,y,z)k^l(z) - \mathbf{U}^{Im}_{\tilde{\beta}_1\ l}(x,y,z)Dk^m_l(z)\Big)\ dz\Big)\tilde{\beta}^J_1(x,y) + \\ \tilde{\beta}^I_1(x,y)\int_\Omega \Big(\mathbf{T}^J_{\tilde{\beta}_1\ l}(x,y,z)k^l(z) - \mathbf{U}^{Jm}_{\tilde{\beta}_1\ l}(x,y,z)Dk^m_l(z)\Big)\ dz\bigg)\,dy\end{aligned}$$

This results in the following expression for $\delta_k\mathcal{J}^2_{\mathrm{AC}}(h)$

$$\delta_k\mathcal{J}^2_{\mathrm{AC}}(h) = \int_\Omega \mathbf{T}^2_{\mathrm{AC}}(x,h)k(x)\,dx - \int_\Omega \mathbf{U}^2_{\mathrm{AC}}(x,h)Dk(x)\,dx,$$

where

$$\mathbf{T}^2_{\mathrm{AC}\ l}(x,h) \ = \ -\frac{1}{\mid\Omega\mid}\int_\Omega\int_\Omega \Theta_{IJ}(y)\Big(\mathbf{T}^I_{\tilde{\beta}_1\ l}(y,z,x)\tilde{\beta}^J_1(y,z) \ + \ \tilde{\beta}^I_1(y,z)\mathbf{T}^J_{\tilde{\beta}_1\ l}(y,z,x)\Big)\,dy\,dz,$$

and

$$\mathbf{U}^{2\ m}_{\mathrm{AC}\ l}(x,h) \ = \ -\frac{1}{\mid\Omega\mid}\int_\Omega\int_\Omega \Theta_{IJ}(y)\Big(\mathbf{U}^{Im}_{\tilde{\beta}_1\ l}(y,z,x)\tilde{\beta}^J_1(y,z) \ + \ \tilde{\beta}^I_1(y,z)\mathbf{U}^{Jm}_{\tilde{\beta}_1\ l}(y,z,x)\Big)\,dy\,dz$$

$\square$

A.4 Computation of $\delta_k \tilde{\Lambda}_{21}(x, h)$

According to equation (15), we have

$$\delta_k \tilde{\Lambda}_{21}(x,h) \quad = \quad \frac{1}{|\,\Omega\,|} \int_\Omega \left(\delta_k \tilde{\beta}_2(x,y,h) \tilde{\beta}_2^T(x,y,h) \; + \; \tilde{\beta}_2(x,y,h) \left(\delta_k \tilde{\beta}_2(x,y,h) \right)^T \right) dy \quad (50)$$

Because of (23), we have

$$\begin{aligned} \delta_k \tilde{\beta}_2(x,y,h) &= \delta_k \Big(\exp\left(-\mathcal{B}(x,h)\right) \beta_2(x,y,h) \Big) \\ &= \Big(\delta_k \exp\left(-\mathcal{B}(x,h)\right) \Big) \beta_2(x,y,h) + \exp\left(-\mathcal{B}(x,h)\right) \Big(\delta_k \beta_2(x,y,h) \Big) \end{aligned}$$

We have already derived, in the previous sections, all we need to evaluate this derivative.

Computation of the first term of $\delta_k \tilde{\beta}_2(x,y,h)$: The first term of $\delta_k \tilde{\beta}_2(x,y,h)$, namely

$$\Big(\delta_k \exp\left(-\mathcal{B}(x,h)\right) \Big) \beta_2(x,y,h),$$

is readily obtained from the derivations carried out in section A.3. Since

$$\delta_k \exp\left(-\mathcal{B}(x,h)\right) = -\mathrm{dexp}\left(-\mathcal{B}(x,h), \delta_k \mathcal{B}(x,h)\right),$$

all the arguments used previously to derive an expression for $\delta_k \exp\left(-\mathcal{A}(x,h)\right)$ are still valid. Replacing ψ by

$$\theta(x,h) = \log\left(\hat{\mu}_1(x)^{-1} \hat{\mu}_2(x,h)\right),$$

results in the expressions

$$\delta_k \theta(x,h) = \int_\Omega \mathbf{t}_\theta(x,z) k(z)\, dz - \int_\Omega \mathbf{u}_\theta(x,z) Dk(z),$$

where

$$\begin{aligned} \mathbf{t}_\theta(x,z) &= \mathrm{dlog}\left(\hat{\mu}_1(x)^{-1} \hat{\mu}_2(x,h), \hat{\mu}_1(x)^{-1} \mathbf{t}(x,z)\right) \\ \mathbf{u}_\theta(x,z) &= \mathrm{dlog}\left(\hat{\mu}_1(x)^{-1} \hat{\mu}_2(x,h), \hat{\mu}_1(x)^{-1} \mathbf{u}(x,z)\right). \end{aligned}$$

Using indexes,

$$\mathbf{t}_{\theta\,\cdot\, l}^{\;\cdot} = \sum_{m=1}^{3} \mathrm{dlog}\left(\hat{\mu}_1(x)^{-1} \hat{\mu}_2(x,h), \left(\hat{\mu}_1(x)^{-1}\right)_{m\,\cdot}\, \mathbf{t}_l^{\cdot\, m}\right)$$

Since the relation between θ and $\mathcal{B}$ is linear ($\mathcal{B} = \mathcal{M}(\theta)$, see equation (24)), we have

$$\delta_k \mathcal{B}(x,h) = \int_\Omega \mathcal{M}(\mathbf{t}_\theta(x,z)) k(z)\, dz - \int_\Omega \mathcal{M}(\mathbf{u}_\theta(x,z)) Dk(z)\, dz,$$

and therefore, using the linearity of $\mathrm{dexp}\left(\cdot,\cdot\right)$ with respect to its second argument:

$$\begin{aligned} &\delta_k \exp\left(-\mathcal{B}(x,h)\right) \beta_2(x,y,h) = \\ &\qquad - \int_\Omega \mathrm{dexp}\left(-\mathcal{B}(x,h), \mathcal{M}(\mathbf{t}_\theta(x,z))\right) \beta_2(x,y,h) k(z)\, dz + \\ &\qquad\qquad \int_\Omega \mathrm{dexp}\left(-\mathcal{B}(x,h), \mathcal{M}(\mathbf{u}_\theta(x,z))\right) \beta_2(x,y,h) Dk(z)\, dz \end{aligned}$$

Computation of the second term of $\delta_k\tilde{\beta}_2(x,y,h)$: The second term of $\delta_k\tilde{\beta}_2(x,y,h)$, namely

$$\exp\left(-\mathcal{B}(x,h)\right)\Big(\delta_k\beta_2(x,y,h)\Big)$$

is readily obtained from the derivations lead in section A.2. Indeed the derivative $\delta_k\beta_2(x,y,h)$ was proved in lemma A.2.1 to be equal to

$$\delta_k\beta_2 = \int_\Omega \mathbf{T}_{1\beta_2}(x,y,z)k(z)\,dz + \mathbf{T}_{2\beta_2}(x,y)k(y) - \int_\Omega \mathbf{U}_{1\beta_2}(x,y,z)Dk(z)\,dz - \mathbf{U}_{2\beta_2}(x,y)Dk(y)$$

Combining these two results allows us to prove the following lemma, analog to lemma A.2.1

Lemma A.4.1. *$\delta_k\tilde{\beta}_2$ can be written as*

$$\delta_k\beta_2(x,y) = \int_\Omega \mathbf{T}_{1\tilde{\beta}_2}(x,y,z)k(z)\,dz + \mathbf{T}_{2\tilde{\beta}_2}(x,y)k(y) - \int_\Omega \mathbf{U}_{1\tilde{\beta}_2}(x,y,z)Dk(z)\,dz - \mathbf{U}_{2\tilde{\beta}_2}(x,y)Dk(y),$$

where the expressions of the tensors $\mathbf{T}_{1\tilde{\beta}_2}$, $\mathbf{T}_{2\tilde{\beta}_2}$, $\mathbf{U}_{1\tilde{\beta}_2}$ and $\mathbf{U}_{2\tilde{\beta}_2}$ are given in the proof.

Proof. We can write immediately

$$\delta_k\tilde{\beta}_2 = \int_\Omega \mathbf{T}_{1\tilde{\beta}_2}(x,y,z,h)k(z)\,dz + \mathbf{T}_{2\tilde{\beta}_2}(x,y,h)k(y) - \int_\Omega \mathbf{U}_{1\tilde{\beta}_2}(x,y,z,h)Dk(z)\,dz - \mathbf{U}_{2\tilde{\beta}_2}(x,y,h)Dk(y),$$

where

$$\begin{aligned}
\mathbf{T}_{1\tilde{\beta}_2}(x,y,z) &= -\mathrm{dexp}\left(-\mathcal{B}(x,h),\mathcal{M}(\mathbf{t}_\theta(x,z))\right)\beta_2(x,y,h) + \exp\left(-\mathcal{B}(x,h)\right)\mathbf{T}_{1\beta_2}(x,y,z,h)\\
\mathbf{T}_{2\tilde{\beta}_2}(x,y) &= \exp\left(-\mathcal{B}(x,h)\right)\mathbf{T}_{2\beta_2}(x,y,h)\\
\mathbf{U}_{1\tilde{\beta}_2}(x,y,z) &= -\mathrm{dexp}\left(-\mathcal{B}(x,h),\mathcal{M}(\mathbf{u}_\theta(x,z))\right)\beta_2(x,y,h) + \exp\left(-\mathcal{B}(x,h)\right)\mathbf{U}_{1\beta_2}(x,y,z,h)\\
\mathbf{U}_{2\tilde{\beta}_2}(x,y) &= \exp\left(-\mathcal{B}(x,h)\right)\mathbf{U}_{2\beta_2}(x,y,h)
\end{aligned}$$

□

This allows us to prove the following

Proposition A.4.1. *The third term, $\delta_k\mathcal{J}^3_{\mathrm{AC}}$, in the expression of $\delta_k\mathcal{J}_{\mathrm{AC}}$ is of the form described in theorem 5.0.1.*

Proof. The proof is completely analog to that of proposition A.2.1. We end up with

$$\delta_k \mathcal{J}^3_{\mathrm{AC}}(h) = \int_\Omega \mathbf{T}^3_{\mathrm{AC}}(x,h)k(x)\,dx - \int_\Omega \mathbf{U}^3_{\mathrm{AC}}(x,h)Dk(x)\,dx,$$

where

$$\begin{aligned}
\mathbf{T}^3_{\mathrm{AC}\ l}(x,h) &= \overline{\widetilde{\mathbf{H}}}(x,h) &&= \int_\Omega \overline{\widetilde{\mathbf{H}}}(z,x,h)\,dz \\
\mathbf{U}^{3\ m}_{\mathrm{AC}\ l}(x,h) &= \overline{\widetilde{\mathbf{K}}}(x,h) &&= \int_\Omega \overline{\widetilde{\mathbf{K}}}(z,x,h)\,dz,
\end{aligned}$$

where the once covariant tensor $\widetilde{\mathbf{H}}(z,x,h)$ is given by the following expression

$$\widetilde{\mathbf{H}}_l(x,y) = \frac{\Phi_{IJ}(x)}{\mid\Omega\mid}\int_\Omega \Big(\mathbf{T}^I_{1\tilde{\beta}_2\ l}(x,z,y)\tilde{\beta}^J_2(x,z) + \tilde{\beta}^I_2(x,z)\mathbf{T}^J_{1\tilde{\beta}_2\ l}(x,z,y)\Big)\,dz + \\ \frac{\Phi_{IJ}(x)}{2\mid\Omega\mid}\Big(\mathbf{T}^I_{2\tilde{\beta}_2\ l}(x,y)\tilde{\beta}^J_2(x,y) + \tilde{\beta}^I_2(x,y)\mathbf{T}^J_{2\tilde{\beta}_2\ l}(x,y)\Big),$$

and the once covariant once contravariant tensor $\widetilde{\mathbf{K}}(z,x,h)$ is given by the following expression

$$\widetilde{\mathbf{K}}^m_l(x,y) = \frac{\Phi_{IJ}(x)}{\mid\Omega\mid}\int_\Omega \Big(\mathbf{U}^{Im}_{1\tilde{\beta}_2\ l}(x,z,y)\tilde{\beta}^J_2(x,z) + \tilde{\beta}^I_2(x,z)\mathbf{U}^{Jm}_{1\tilde{\beta}_2\ l}(x,z,y)\Big)\,dz + \\ \frac{\Phi_{IJ}(x)}{2\mid\Omega\mid}\Big(\mathbf{U}^{Im}_{2\tilde{\beta}_2\ l}(x,y)\tilde{\beta}^J_2(x,y) + \tilde{\beta}^I_2(x,y)\mathbf{U}^{Jm}_{2\tilde{\beta}_2\ l}(x,y)\Big).$$

□

References

[1] D.C. Alexander and J.C. Gee. Elastic matching of diffusion tensor images. *Computer Vision and Image Understanding*, 77:233–250, 2000.

[2] D.C. Alexander, J.C. Gee, and R. Bajcsy. Similarity measures for matching diffusion tensor images. In *British Machine Vision Conference*, page 251, 1999.

[3] D.C. Alexander, J.C. Gee, and R. Bajcsy. Strategies for data reorientation during non-rigid warps of diffusion tensor images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 463–472, 1999.

[4] V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A log-euclidean framework for statistics on diffeomorphisms. In *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 924–931, 2006.

[5] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56:411–421, 2006.

[6] Vincent Arsigny. *Processing Data in Lie Groups: An Algebraic Approach. Application to Non-Linear Registration and Diffusion Tensor MRI*. PhD thesis, Ecole Polytechnique, 2006.

[7] H. Brezis. *Analyse fonctionnelle. Théorie et applications*. Masson, 1983.

[8] Y. Cao, M.I. Miller, S. Mori, R.L. Winslow, and L. Younes. Diffeomorphic matching of diffusion tensor images. In *Conference on Computer Vision and Pattern Recognition Workshop*, page 67, 2006.

[9] P.G. Ciarlet. *Mathematical Elasticity*, volume 1. North Holland, 1988.

[10] R. Courant. *Calculus of Variations*. New York, 1946.

[11] L.C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Proceedings of the American Mathematical Society, 1998.

[12] O. Faugeras and G. Hermosillo. Well-posedness of two non-rigid multimodal image registration methods. *Siam Journal of Applied Mathematics*, 64(5):1550–1587, 2004.

[13] P.T. Fletcher and S. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proc. Computer Vision Approaches to Medical Image Analysis*, pages 87–98, Prague, May 2004.

[14] K. Grove, H. Karcher, and E.A. Ruh. Group actions and curvature. *Inventiones Mathematicae*, 23:31–48, 1974.

[15] A. Guimond, C.R.G. Guttmann, S.K. Warfield, and C.F. Westin. Deformable registration of DT-MRI data based on transformation invariant tensor characteristics. In *IEEE International Symposium on Biomedical Imaging*, Washington (DC), USA, July 7–10 2002.

[16] A. Guimond, A. Roche, N. Ayache, and J. Meunier. Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Transactions in Medical Imaging*, 20(1):58–69, 2001.

[17] Gerardo Hermosillo. *Variational Methods for Multimodal Image Matching*. PhD thesis, Université de Nice-Sophia Antipolis, May 2002.

[18] D. K. Jones, L.D. Griffin, D.C. Alexander, M. Catani, M.A. Horsfield, R. Howard, and S.C. Williams. Spatial normalization and averaging of diffusion tensor MRI data sets. *Neuroimage*, 17(2):592–617, 2002.

[19] Alexander Leemans. *Modeling and Processing of Diffusion Tensor Mangetic Resonance Images for Improved Analysis of Brain Connectivity*. PhD thesis, University of Antwerp, Belgium, 2006.

[20] C. Lenglet, M. Rousson, R. Deriche, and 0. Faugeras. Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing. *Journal of Mathematical Imaging and Vision*, 25(3):423–444, October 2006.

[21] Christophe Lenglet. *Geometric and Variational Methods for Diffusion Tensor MRI Processing*. PhD thesis, Université de Nice Sophia-Antipolis, 2006.

[22] R. Mneimné and F. Testard. *Introduction à la théorie des groupes de Lie classiques*. Hermann, 1986.

[23] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):735–747, April 2005.

[24] I. Najfeld and T. Havel. Derivatives of the matrix exponential and their computation. Technical Report TR-33-94, Harvard Medical School, Department of Biological Chemistry and Molecular Pharmacology, November 1994.

[25] H.J. Park, M. Kubicki, M.E. Shenton, A. Guimond, R.W. McCarley, S.E. Maier, R. Kikinis, F.A. Jolesz, and C.F. Westin. Spatial normalization of diffusion tensor MRI using multiple channels. *Neuroimage*, 20(4):1995–2009, 2003.

[26] X. Pennec. Left-invariant riemannian elasticity: A distance on shape diffeomorphisms? In *Proceedings of International Workshop on the Mathematical Foundations of Computational Anatomy*, pages 1–13, 2006.

[27] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

[28] G.K. Rohde, S. Pajevic, and C. Pierpaoli. Multi-channel registration of diffusion tensor images using directional information. In *IEEE International Symposium on Biomedical Imaging*, pages 712–715, Washington (DC), USA, April 15–18 2004.

[29] J. Ruiz-Alzola, C.F. Westin, S.K. Warfield, C. Alberola, S.E. Maier, and R. Kikinis. Nonrigid registration of 3d tensor medical data. *Medical Image Analysis*, 6(2):143–161, 2002.

[30] J. Ruiz-Alzola, C.F. Westin, S.K. Warfield, A. Nabavi, and R. Kikinis. Nonrigid registration of 3d scalar, vector and tensor medical data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 541–550, 2000.

[31] Raimundo Sierra. Nonrigid registration of diffusion tensor images. Master's thesis, Swiss Federal Institute of Technology (ETHZ), March 2001.

[32] L.T. Skovgaard. A Riemannian geometry of the multivariate normal model. Technical Report 81/3, Statistical Research Unit, Danish Medical Research Council, Danish Social Science Research Council, 1981.

[33] H. Zhang, P.A. Yushkevich, and J.C. Gee. Towards diffusion profile image registration. In *IEEE International Symposium on Biomedical Imaging*, pages 324–327, Washington (DC), USA, April 15–18 2004.

[34] H. Zhang, P.A. Yushkevich, and J.C. Gee. Deformable registration of diffusion tensor MR images with explicit orientation optimization. In *International Conference on Medical Image Computing and Computer Assisted Intervention*, pages 172–179, 2005.

[35] Hui Zhang, Paul A. Yushkevich, and James C. Gee. Registration of diffusion tensor images. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 842–847, 2004.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399